National Defence / Défense nationale

UNCLASSIFIED

DRES

SUFFIELD REPORT

NO. 577

AD-A259 371

MOTION AND STRUCTURE ESTIMATION OF MANOEUVRING OBJECTS IN MULTIPLE-CAMERA IMAGE SEQUENCES

DTIC
ELECTE
JAN 21 1993
S C D

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

93-01002
14608

by

Victor C. Aitken

November 1992

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD : RALSTON : ALBERTA

WARNING
"The use of this information is permitted subject to recognition of proprietary and patent rights".

Canada

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD
RALSTON, ALBERTA

SUFFIELD REPORT NO. 577

MOTION AND STRUCTURE ESTIMATION OF MANOEUVRING OBJECTS IN MULTIPLE-CAMERA IMAGE SEQUENCES

by

Victor C. Aitken

DTIC QUALITY INSPECTED 5

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

WARNING
"The use of this information is permitted subject to recognition of proprietary and patent rights".

Abstract

Estimation of structure and six-degree-of-freedom motion of manoeuvring objects through measurements of feature positions in long, multiple-camera image sequences is widely recognized to have broad industrial, military, and space applications, particularly in the control of autonomous systems. This report focuses on the case of manoeuvring objects thereby removing restrictive assumptions concerning the mode of translational and rotational motion which are commonly employed in many existing methods. Object manoeuvres, being "smooth" and time correlated, are modelled as first-order Gauss-Markov processes for both translational and rotational motion. In the literature, rotational motion is often parameterized with unit quaternions even though constraints on the quaternion norm are not easily enforced, roll-pitch-yaw parameterizations have been reported to be poorly behaved and have led to computationally demanding implementations, and results using the Euler angle-axis parameterization in recursive motion and structure estimation are not available. In this report, six-degree-of-freedom, nonlinear, approximate state estimation filters for quaternion, roll-pitch-yaw, and angle-axis parameterizations are compared in terms of estimation performance for manoeuvring object trajectories. Special consideration is given to the problem of imposing unit norm on the estimated quaternion since previously proposed methods led to filter instability, particularly in angular velocity estimation. Methods proposed herein not only demonstrate how previous quaternion-based algorithms might be extended to track manoeuvring objects observed in multiple-camera image sequences, but also provide two much simpler alternatives using angle-axis and roll-pitch-yaw parameterizations. The angle-axis filter was found to give the best overall performance with a significant reduction in computational requirements in comparison to the quaternion filter.

Executive Summary

Image-based motion analysis is primarily concerned with accumulation and refinement of information related to motion and structure of objects within the environment of an imaging system. Motion parameters include position, orientation, and time derivatives of position and orientation of objects with respect to a specified reference frame. Structure, in the present context, refers to the shape of observed objects in the sense that positions of object features are estimated with respect to an object-centred reference frame. Estimation of structure and six-degree-of-freedom motion of manoeuvring objects through measurements of feature positions in long, multiple-camera image sequences is widely recognized to have broad industrial, military, and space applications. Image sequence analysis provides a means for passive tracking of either airborne or ground-based moving targets. Structure estimation in a static environment and recovery of motion of a moving platform which carries the imaging system is an equivalent problem which is of fundamental importance in the control and guidance of autonomous systems.

The Defence Technologies Division of the Defence Research Establishment Suffield has several groups whose research efforts are or will be focussed in part on acquiring information from imaging systems: object localization, tracking and recognition is of primary interest to the Threat Detection Group; methods for passive target tracking are being investigated by the Systems Integration and Vehicle Concepts Groups; and vision-based navigation and control of autonomous systems are common research areas of the Advanced Guidance Concepts, Electronic Design and Instrumentation, and Vehicle Concepts Groups.

This report provides a framework on which future research in motion and structure estimation can be based. Methods proposed herein concentrate primarily on the recovery of motion and structure of a rigid object observed with a multiple-camera imaging system. However, the resulting state estimation algorithms are sufficiently general to be of interest in a broad range of applications, including those of the various research groups listed above. The report provides a detailed review of image analysis methods currently available in the literature and proposes a hierarchical image analysis system which partitions the overall problem into three coupled target tracking problems with increasing levels of complexity. One of these levels is concerned with the estimation of structure and six-degree-of-freedom motion of observed objects and is the focus of the remainder of the report. Detailed background material is provided in the Appendices in order that this report may be self-contained.

A fundamental difficulty associated with vision-based motion estimation, in comparison to the more traditional radar-based point tracking problem for example, lies in the recovery of the object's orientation and rotational motion. Theoretically precise models of rotational motion may be formulated in terms of the object's inertial

parameters, external torques acting on the object, and Euler's equations which represent a coupled system of three nonlinear differential equations that admit closed-form solutions only in very limited cases. In object tracking applications, however, inertial parameters and external torques are generally not available. Consequently, alternate and often simplified models for rotational motion are employed.

Many existing methods for recursive motion and structure estimation from long image sequences have been developed under specific assumptions concerning the nature of translational and rotational motion, for example, constant velocity, acceleration, precession, or angular momentum. A number of techniques define object structure based in part on the assumed motion of the object which can lead to unobservability for some manoeuvres. Rotational motion is often modelled with unit quaternions for which dynamical models are parameterized directly in terms of angular velocity and are nonlinear. Roll-pitch-yaw parameterizations have been reported to be poorly behaved numerically and have led to computationally demanding implementations. Published results using the Euler angle-axis parameterization in recursive motion and structure estimation are not present in the literature.

Methods proposed herein focus in part on removing restrictive assumptions concerning the nature of object motion and eliminating dependence of object structure on assumed motion by developing motion, structure, and measurement models for a rigid manoeuvring object observed with a multiple-camera imaging system. Object manoeuvres, being "smooth" and time correlated, are modelled as first-order Gauss-Markov processes for both translational and rotational motion. Object structure is defined on the basis of observed feature points only. Six degree-of-freedom extended Kalman filters for roll-pitch-yaw, Euler angle-axis, and quaternion parameterizations are developed and compared in terms of estimation performance for manoeuvring object trajectories.

Both the angle-axis and roll-pitch-yaw filters are based on approximate linear dynamic models which lead to computationally efficient implementations. Computational requirements of the quaternion filter are more significant than the other two due primarily to the nonlinear dynamic model. Another difficulty associated with the quaternion filter is that nonlinear constraints which maintain unit quaternion norm are not easily incorporated into the linear structure of the Kalman filter. Previously proposed methods of applying an impulsive normalization of the estimated quaternion immediately following each observation event was found to contribute significantly to divergence in angular velocity estimates. Instead, the quaternion estimate is propagated without impulsive normalization, and a unit quaternion as well as appropriately scaled structure vectors are extracted from the state estimate.

Simulation results with computer-generated imagery suggest that all three filters perform equally well for most trajectories. The roll-pitch-yaw filter has the simplest

form, but sometimes yields poor results in angular velocity estimates. The quaternion filter shows degraded performance in estimation of object position and some structural parameters. The angle-axis filter is computationally efficient and performs at least as well as, and often better than, the quaternion and roll-pitch-yaw filters. This work not only demonstrates how previously proposed quaternion-based approaches might be extended to track manoeuvring objects observed in multiple-camera image sequences, but also provides two much simpler alternatives using angle-axis and roll-pitch-yaw parameterizations. The angle-axis filter based on a linear dynamic model was found to give the best overall performance with a significant reduction in computational requirements in comparison to the quaternion filter.

Future research efforts will address problems associated with implementation of other modules of the hierarchical structure proposed in Section 1.2. Generally, the proposed system is composed of three coupled multi-target tracking systems with increasing levels of complexity. It might be expected that statistical-based multi-target tracking methods, which have been extensively studied in radar-based point tracking applications, will be extremely useful in such a system. Feature detection and correspondence as well as feature occlusion processes represent challenging problems for future investigations. However, a primary consideration in future work will also involve further evaluation of the extended Kalman filtering approach in comparison to other nonlinear observers. Adaptive extended Kalman filtering, Lyapunov methods, transformation to nonlinear observer canonical form, nonlinear map inversion, and sliding mode observers are representative of recently proposed methods which may lead to more robust observers for recovery of object motion and structure from multiple-camera image sequences. Initially, future work will focus on performance evaluations with simulated imagery. As motion and structure estimation algorithms mature, however, experiments with real imagery will be required to assess the true performance of proposed image analysis systems.

Table of Contents

List of Figures

List of Tables

1. Introduction

A primary function of biological vision systems is the detection, processing, and understanding of motion perceived from a continuous flow of visual information. Perceived motion results from relative motion of a vision system within a dynamic environment. Analysis of object motion in scenes, or the apparent motion perceived by a moving observer in a static scene can yield vital information concerning the structure and dynamic behavior of the observed process. Over the last two decades, research in motion detection, analysis, and understanding techniques with applications in computer vision systems for robotics and autonomous vehicles represents an area of vigorous growth. Rosenfeld's annual bibliographical surveys (see [1] for example) demonstrate the intense interest in this field of research.

Motion and structure estimation from multiple-camera image sequences has broad industrial, military, and space applications. Motion parameters include position, orientation, and time derivatives of position and orientation of an object with respect to a specified reference frame. Structure, in the present context, refers to the shape of the observed object in the sense that feature positions are estimated with respect to an object-centred reference frame. These notions are made more precise in Chapters 2 and 3. A fundamental difficulty associated with vision-based motion estimation, in comparison to the more traditional radar-based point tracking problem for example, lies in the recovery of object orientation and rotational motion. Theoretically precise models of rotational motion may be formulated in terms of the object's inertial parameters, external torques acting on the object, and Euler's equations which represent a coupled system of three nonlinear differential equations that admit closed-form solutions only in very limited cases. In object tracking applications, inertial parameters and external torques are generally not available. Consequently, alternate and often simplified models for rotational motion are employed.

Many existing approaches to motion and structure recovery from long image sequences are somewhat restrictive in that solutions are proposed for specific cases of translational and rotational motion, for example, constant velocity, acceleration, precession, or angular momentum, etc. These techniques have been designed and evaluated on appropriate data generated under these assumptions. However, such precise *a priori* information is seldom available in general object tracking applications, particularly in an unstructured environment. Rotational motion is often parameterized with

unit quaternions, which leads to further difficulties when enforcing a nonlinear constraint on the quaternion norm. Roll-pitch-yaw parameterizations, on the other hand, have been reported to be neither simple nor well-behaved numerically and have led to overly complicated and computationally demanding implementations. The Euler angle-axis parameterization has appeared in small intersample-angle approximations, but published results in the context of recursive motion and structure estimation are not available. None of these methods have been developed for the case of manoeuvring objects with measurements from multiple-camera systems. These observations are discussed further in the literature review of Section 1.1. The reader who might be unfamiliar with these parameterizations, after becoming familiar with notation introduced in Chapter 2, may wish to consult Appendix A which presents a brief tutorial review.

Primary contributions of the research reported herein include the development of motion, structure, and measurement models, and corresponding recursive nonlinear filters to estimate motion and structure of a manoeuvring object through measurements taken from a multiple-camera imaging system. Six degree-of-freedom, nonlinear, approximate state estimation filters for three parameterizations of rotational motion, using roll-pitch-yaw, Euler angle-axis, and quaternion parameters, are developed and compared in terms of estimation performance for manoeuvring object trajectories. Simulation studies have been conducted over a wide range of trajectories. Simulation results suggest that the roll-pitch-yaw, Euler angle-axis, and quaternion filters perform equally well in almost all cases, with the roll-pitch-yaw filter giving slightly poorer results and the angle-axis filter giving slight improvements over the the other two. The roll-pitch-yaw filter has the simplest form due to the use of a linear dynamic model. The Euler angle-axis filter is also derived from a linear dynamic model which results in computationally efficient implementation, but requires an occasional impulsive reset of the state estimate to maintain the rotation angle to within $\pm\pi$. The quaternion filter is the most computationally demanding of the three due to a nonlinear dynamic model. Time propagation of the estimated quaternion is performed with a simple numerical integration scheme which incorporates estimates of higher-order time derivatives of angular velocity. Special consideration is given to the problem of imposing unit norm on the estimated quaternion, since previously proposed methods led to filter instability, particularly in angular velocity estimation. Strict assumptions are not imposed on the mode of translational or rotational motion except "smoothness" conditions in the sense that parameters may be differentiated with respect to time.

This introductory chapter continues with a discussion of motion analysis techniques through a review of the literature. Section 1.1 describes and compares two fundamentally different approaches to motion analysis known as image flow-based and feature-based techniques, and presents a summary of major contributions to motion analysis which are based on extended Kalman filtering. Section 1.2 discusses various components of feature-based motion analysis and proposes a possible hierarchical

structure for an image analysis system. Three-dimensional, six degree-of-freedom motion and structure estimation, which is the primary subject of this investigation, represents only a single component of the hierarchical system. Section 1.3 summarizes this introductory chapter and provides an overview of the remainder of the report.

1.1 Literature Review

Early work in image sequence analysis prior to 1978 has been surveyed by Martin and Aggarwal [2] and Nagel [3]. A large amount of research during this period focussed on applications of motion analysis to bandwidth compression for transmission of TV signals, and on detection and tracking of geographical landmarks and cloud motion in satellite imagery. In a more limited sense, real-time feature tracking techniques were also studied in applications of visual feedback to industrial automation and autonomous vehicle navigation. Feature tracking methods initially focussed on extracting motion parameters through detection of object displacement between two consecutive frames—a technique that is still being investigated today. Some emphasis, however, shifted towards tracking features within a sequence of more than two image frames through stochastic modeling of the dynamic processes being observed.

Motion analysis techniques are often classified according to whether they are formulated through an image flow-based, or feature-based approach. The image flow approach first estimates apparent velocities of intensity patterns in an image sequence before motion analysis begins. In feature-based algorithms, features such as points, lines, contours, corners, etc., are extracted from each image and matched between images in order to estimate relative motion of features in scenes with respect to an observer. More recent surveys have been presented by Nagel [4], who focuses on techniques based on measurement of image flow, and Aggarwal [5], who treats primarily feature-based techniques. Feature matching between successive images is known as the *correspondence problem* and has been reviewed by Aggarwal, Davis, and Martin [6]. The correspondence problem, although it has not yet been explicitly treated in this research, is discussed further in Section 1.2 in terms of its placement in the overall motion analysis problem.

Image flow is often modeled by the image flow constraint equation [8] which describes the interaction between velocity fields and spatial and temporal gradients of image intensity functions under a common assumption of small motion between successive image frames. This approach requires that spatially and temporally discrete intensity distributions be differentiated through approximate differencing techniques which tends to enhance noise present in real images. The image velocity field is then iteratively computed for each pixel based on various proposed optimization criteria [8, 9, 13]. The image flow field contains information concerning relative depth, surface orientation, and relative motion of objects [10]. Several methods have been proposed to extract three dimensional structure and motion from flow fields [11, 12, 14]. It

has been noted, however, that the system of equations relating position and motion in space to image position and flow are not well behaved numerically and are highly sensitive to noise in image position and flow estimates [5]. Further difficulties are encountered in the accurate measurement of flow fields near discontinuities in image intensity and in areas of uniform intensity [8]. Moreover, if the inter-frame motion is indeed small, inaccuracies in flow field estimates may overshadow the true velocity field and lead to ambiguous results.

Feature-based approaches formulate systems of equations relating positions of features detected in an image sequence to object translational and rotational parameters. This approach implicitly assumes that the object of interest has a set of discernible features which can be detected in the image sequences. Roach and Aggarwall [15] give a thorough discussion of this problem when only two or three frames are considered. They proposed a modified least-squared error solution of a nonlinear system of equations to estimate motion parameters. Their results demonstrate that this approach can be quite sensitive to noise in estimates of feature positions in the image. Tsai and Huang [16, 17] have shown that in many cases motion parameters can be obtained by solving a system of linear equations for a set of "essential parameters", followed by a Singular Value Decomposition of the 3×3 essential parameter matrix. Their simulation results also show that motion parameter estimates can be highly sensitive to feature position errors. Fang and Huang [18] have implemented the approach and give detailed experimental results. Again, accuracy of the results are seriously degraded in the presence of feature position measurement errors. Recently, extensions of this method, such as those proposed by Philip [22] or Weng *et al* [19, 20, 21] for example, aim at reducing measurement noise sensitivity through refinement of parameter estimates over multiple frames.

Structure and motion processes, and the correspondence process are believed to be complementary, perhaps even inseparable, in the human visual system [73]. Feature detection and correspondence are required to interpret raw image data and drive the motion estimation scheme. Knowledge of motion parameters and their uncertainty, on the other hand, may serve as constraints to reduce the search space associated with detection and correspondence problems. Matthies and Kanade [43] give a thorough discussion of this concept and provide two illustrative examples which propose solutions based on Kalman filtering techniques. The Kalman filter with its recursive structure and explicit modeling of uncertainty has been proposed as an extremely useful approach to dynamic vision problems, particularly in view of the "built in" state and error prediction capabilities. Applications of Kalman filtering to image analysis have been reported by many authors, for example Futrelle and Speckert [23], Gennery [24], Hallman [25], Legters and Young [26], Stuller and Krishnamurthy [27], Wu *et al.* [38], Matthies, Kanade, and Szeliski [44], Chang *et al.* [39], and Iu and Wohn [53, 54] consider diverse vision-based problems to which sequential state estimation has been applied. The work of three other groups, however, provides the primary motivation for the approach investigated in this report.

Ayache and Faugeras [40], Ayache and Lustman [41], Randall, Foret, and Ayache [42] (see further references cited therein) recursively build and update three dimensional representations of the environment of a mobile robot. These methods are developed primarily for trinocular (three cameras) image sequences, however their general approach could also be applied to binocular stereo imagery or that produced by a monocular imaging system under egomotion. Positions of world features (points, lines, planes) are recursively estimated from measurements in the image plane through extended Kalman filtering. They have generalized the problem of updating positions of any feature so that the same Kalman filtering structure can be employed in all cases. Geometric relations are inferred between features by computing a generalized Mahalanohis distance and using a χ^2 acceptance test. An estimate of robot motion is also maintained with the Kalman filter. They have applied their technique to real scenes taken from an indoor mobile platform and demonstrate some of the major properties of their algorithm.

Dickmanns and Graefe [45] present a thorough discussion of their approach to dynamic machine vision which employs special hardware and methods for feature extraction and information processing. Through the use of integral spatio-temporal models, extended Kalman filtering, and prediction error feedback, they have demonstrated real-time vision-based control capabilities in several important problems in robotics [46]. Real-time capability in dynamic vision is achieved through a specialized multiple instruction stream, multiple data stream (MIMD) computer architecture [46, 47]. This system associates each processing element with a dynamically allocated window on each image which, in turn, is associated with a single feature of interest. Each window is defined by prediction error feedback from the Kalman filter. Controlled correlation techniques are employed by each processing element to efficiently extract the feature of interest in each window. Results of feature extraction are passed from all processing elements to a host computer which executes the extended Kalman filter. They have successfully applied their technique to control of an inverted pendulum on a moving cart [48], autonomous vehicle docking problems [46], autonomous road vehicle guidance [46], and vision-based control of an aircraft in a simulated landing approach [49]. More recently, enhancement of road curvature models and extensions of algorithms and hardware to obstacle avoidance in autonomous vehicle guidance applications have been proposed [50, 51, 52].

Broida [37], Broida and Chellappa [28, 29, 30, 31, 36], Broida, Chandrashekhar, and Chellappa [32], and Young and Chellappa [33, 34, 35] have applied extended Kalman filtering techniques to estimation of motion and structural parameters of a rigid object through measurements taken from noisy monocular and stereo image sequences. Early work, [28], considered a special case of the general motion problem in which object translational motion is constrained to lie on a plane defined by the optical axis and central scan line of the imaging system. All feature points of interest are assumed to lie on this plane. Estimates of model parameters were refined over an arbitrary number of image frames with an iterated extended Kalman filter under

the assumption that some prior knowledge of object structure is available. In three dimensional tracking, [29]-[32], rotational motion is modelled with unit quaternions. An earlier choice of Euler angles and rates was described as being neither simple nor well-behaved, which is in sharp contrast to the results reported herein. Several cases have been treated based on *a priori* information concerning the nature of object motion: pure translational motion, [29]; constant translational and rotational velocities, [32, 36]; and constant translational acceleration and rotational motion with constant precession, [33, 34]. Both recursive algorithms, [28, 29, 32, 33, 34], and a batch approach based on a fixed number of frames, [30, 36], have been investigated. Approximate Cramer-Rao covariance bounds in the absence of *a priori* information are used extensively for performance analysis, [31]. Considerable effort has also been devoted to uniqueness issues for the constant acceleration and precession case, [34]. A more detailed review of their recursive methods for monocular systems, [28, 32], is given by Aitken [58].

Methods reviewed above have been developed for significantly diverse applications. Each method is dependent on a particular set of assumptions. For example, different types of imaging sensors have been investigated: sonar [25], visual monocular [23] [26]-[32] [38, 39, 44, 45, 50], or visual stereo and trinocular [24, 33, 34, 40, 41, 45, 50]. Some approaches assume a static environment with moving sensor [25, 40, 41, 44], while others consider dynamic environments with stationary sensors [23, 24, 26] [28]-[39], and some consider both a dynamic environment and moving sensor [27, 45, 50]. In some cases only motion parallel to the image plane is considered [26, 27, 44]. Of those techniques which estimate motion and orientation of objects in scenes, some assume partial or complete knowledge of object structure [23, 24, 28, 38] while others estimate total object structure from stereo images or with monocular image sequences assuming depth information is available [32, 33, 34]. A general result is that performance of the above techniques cannot be directly compared since each must be applied to the specific problem for which it was developed.

Evaluation of proposed estimation schemes for long image sequences using real imagery has been conducted extensively only in the work of Ayache and Faugeras [40], Matthies *et al.* [44], and Dickmans *et al.* [46]. In a more limited sense, Wu *et al.* [38] show results for imagery obtained in a controlled laboratory setting, while Broida [37] and Broida *et al.* [32, 36] show qualitative results based on real imagery for which ground truth data is not available, and feature points were extracted manually. In most of the above investigations, simulations are performed in which feature detection and correspondence problems are assumed to be solved. In this case, feature positions in each image are obtained from true analytical motion and structural models and an assumed image formation model, and additive noise processes are included to simulate measurement noise.

Application of sequential state estimation techniques such as the Kalman filter to machine vision problems emphasizes temporal continuity conditions for image sequence interpretation. This allows proper definitions of state variables and formula-

tion of systems of differential equations to model motion processes of objects in space and time. Dickmanns [45] refers to spatio-temporal motion analysis as "dynamic vision" which, in contrast to "static" image sequence processing, does not separately apply object recognition from one frame to the next as a first step and motion reconstruction afterwards as a second. Instead, object structure and motion are treated simultaneously in order to generate an estimate of the system state based on noisy measurements from images. There are several immediate advantages of this approach:

1. Storage requirements for past images are reduced or eliminated in that the current system state can be represented solely by state vectors and their uncertainty, together with shape descriptors which describe the particular feature to which each state vector applies (point, line, plane, etc.);

2. The motion estimation process need not be explicitly dependent on assumptions of small motion between scenes which is common to the optical flow based approach [8, 9, 11, 12, 13];

3. The estimation process can be made robust to large measurement errors through explicit modeling of measurement uncertainty; and

4. Feedback of state predictions and uncertainties reduces the search space associated with feature detection and correspondence problems.

Two disadvantages of this technique arise from computational requirements of Kalman filters, and nonlinear dynamical and measurement models which result in the use of suboptimal extended Kalman filters. The computational burden may, in some cases, be less than that required for iterative approaches to obtain solutions to systems of nonlinear equations. The use of suboptimal extended Kalman filters in computer vision and other nonlinear observer applications is often criticized for several reasons:

1. Guaranteed convergence or stability cannot be established *a priori*;

2. Dynamical models are often approximated as polynomials in time which may contribute to modelling error and instability; and

3. Theoretical requirements of precise stochastic models, such as white Gaussian disturbances with known mean and covariance, are often violated.

The fact remains, however, that extended Kalman filters are being used in a wide range of applications with varying degrees of success. In these applications, extended Kalman filters should be considered as nonlinear observers with a particular structure in which the plant-, and in some cases the measurement-, "covariance" matrices are treated as tuning parameters which are specified by the designer in order to achieve a desired performance over a sufficient range of trajectories.

In planar motion applications, the axis of rotation is fixed and is often assumed to be known [15, 25, 26, 28, 58, 59]. Parameterization in terms of angular velocity then results in linear dynamical models. More general rotational motion is often modelled with unit quaternions [24, 53, 55], [29]-[37]. Dynamical models which govern time propagation of the unit quaternion are parameterized in terms of angular velocity and are nonlinear. State estimation is then subject to a nonlinear constraint which imposes unit length on the four-element quaternion. This constraint is not easily incorporated into the linear structure of the Kalman filter. A common approach, [53], [29]-[37], is to normalize the estimated quaternion following each measurement update of the filter. However, in the present investigation with manoeuvring objects, this practice was found to contribute significantly to divergence in angular velocity estimates. Bar-Itzhack and Oshman [65] also observed poor performance with impulsive normalization even when angular velocity measurements are available. An approach similar to that employed by Bar-Itzhack and Oshman, [65], is used in that the quaternion estimate is propagated in the filter equations without impulsive normalization. For output purposes only, a unit quaternion as well as object structure vectors are extracted from the state estimate. This problem is discussed further in Section 3.5.2. The use of higher-order time derivatives of angular velocity in a quaternion-based filter has been examined by Iu and Wohn [53] who employ a dynamical model for monocular image sequences similar to Broida *et al.* [32, 36]. However, time propagation in their approach assumes constant angular velocity over the sample period. Methods proposed herein assume instead that the highest time derivative of angular velocity included in the state vector is approximately constant over any sample period. A numerical integration scheme which maintains constant quaternion norm is derived for time propagation of the quaternion estimate.

Quaternions are often selected over roll-pitch-yaw angles due to mathematical singularities of the roll-pitch-yaw parameterization. Wu *et al.*, [38] use roll-pitch-yaw angles and parameterize rotational motion in terms of a constant inter-frame rotation matrix. This approach leads to an extremely complicated and computationally demanding implementation. Their technique has been applied to real scenes generated in a laboratory environment using 14 feature points on an object of known structure which is constrained to lie on a planar surface. Previous research, [58], has demonstrated that very little additional information is provided to the planar motion problem as the number of feature points is increased beyond 4. In the present investigation, both the roll-pitch-yaw and Euler angle-axis filters are developed from approximate linear dynamical models. In this case rotational motion is not parameterized directly in terms of angular velocity and its time derivatives. Instead, angular velocity is computed through a nonlinear function of the state estimates and compared to the true value and that generated by the quaternion filter.

This Section has provided a review of only a small subset of literature available on motion and structure estimation and has focussed on those methods which employ extended Kalman filtering. Three dimensional motion and structure estimation repre-

sents only a single component of the overall motion analysis problem. The following Section discusses a possible structure for a motion analysis system and describes some of the assumptions, restrictions, and limitations of the methods studied in the remainder of this report.

1.2 Hierarchical Motion Analysis

The motion analysis problem is primarily concerned with accumulation and refinement of information pertaining to the motion and/or structure of objects within the environment of an imaging system. One possible hierarchical approach to feature-based motion analysis is shown in Figure 1.1. The system has been divided into eight levels, as indicated on the right, in which higher levels represent increased sophistication in the representation of observed processes. In Level I (bottom row of the Figure), multiple cameras, which need not be stationary, observe a dynamic scene. An estimate of each camera's motion is computed from local sensors and perhaps feedback from higher levels of the system. In Level II, each camera produces an image sequence which must be captured and digitized, and may pass through a pre-processing stage. Feature detection algorithms are then applied in Level III to each frame, or a subsequence of frames, in order to extract features of interest (points, lines, corners, etc.). Feedback from motion estimation algorithms in higher levels of the system may significantly reduce the computational effort required for feature detection. It may also be necessary to scan any remaining portions of image frames for new features, or features that might reappear after being occluded. The output data from Level III includes sets of available features from each image sequence.

Level IV of the system attempts to establish correspondences between features extracted from image sequences. In a temporal sequence of images generated by any single camera, the process of tracking features from one sample time to the next will be referred to as "temporal correspondence." In the case of feature points, this is essentially a two-dimensional multi-target tracking problem. Salari and Sethi [57] have recently proposed a batch approach for temporal correspondence in the presence of occlusion. This method is an extension of that proposed by Sethi and Jain [56]. However, their approach is based on the assumption that, "As long as these points are on physical objects, we can safely assume that the trajectories of these points in the stationary image plane will be smooth." This assumption is not correct in the context of their definition of "smooth trajectories". Smooth object motion can and often does result in sharp cusps in image plane trajectories (see Figure 2.3 of this report) which violates their smoothness definition. If feature point trajectories are modelled with parametric equations in time $(u(t), v(t))$, "smooth" three-dimensional object motion will result in $u(t)$ and $v(t)$ being "smooth" functions of time. However, if the image plane trajectory is written as $v(t) = f[u(t)]$, it is very restrictive to assume, as in [56], that $f[\cdot]$ is smooth. A possible solution to the temporal correspondence problem may be provided by sequential state estimation

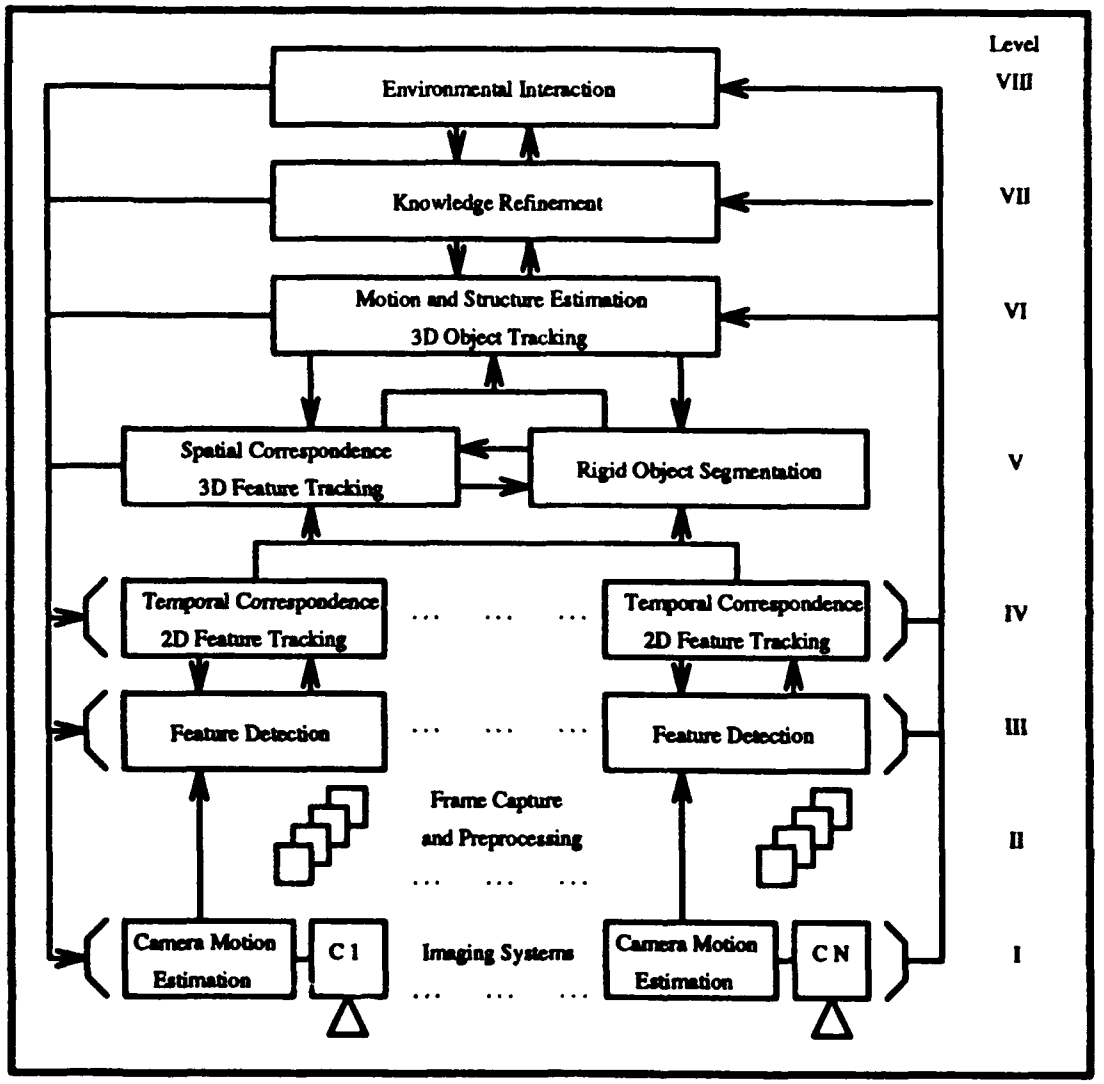

Figure 1.1

Hierarchical structure for motion analysis.

techniques, such as the Kalman filter. This approach has received a great deal of attention in radar-based multi-target tracking literature in which adaptive Kalman filtering or hypothesis testing methods are proposed to track multiple manoeuvring targets. Again, feedback from motion and structure components of higher levels of the system may guide the two-dimensional tracking, or temporal correspondence problem.

Once feature positions and temporal correspondence have been established for image sequences of each camera, higher levels of the system attempt to fuse information from multiple cameras. Level V focuses on three-dimensional feature tracking and segmentation of feature sets into groups such that each group corresponds to a rigid object. The process of matching features seen in two or more images taken at the same time instant will be referred to as "spatial correspondence." Three dimensional information can be inferred from image plane feature position measurements and spatial correspondence in two or more cameras. As a result, this processing stage leads to a three-dimensional multi-target tracking problem.

Initial spatial correspondences must be established through epipolar geometry together with any information available from two-dimensional tracking problems of Level IV. Spatial correspondence for binocular stereo pairs is simplified if the image planes are parallel and physically close together. In this case corresponding points have almost the same locations and possibly similar appearances in the respective images. However, triangulation then becomes ill-conditioned and highly sensitive to measurement noise due to the nearly parallel three dimensional rays representing inverse perspective projections of the corresponding points. Spatial correspondence is simplified and more robust if more than two cameras are employed [41].

With image plane feature position measurements, and temporal and spatial correspondence, sequential state estimation, possibly in the form of adaptive Kalman filtering with hypothesis testing, may provide a means to track features in three dimensions. Three dimensional position and velocity information will be useful to segment all available features into rigid objects. For example, in the case of feature points, if vectors joining any two three-dimensional points have constant length (in a statistical sense, i.e. through hypothesis testing), then associate these two points with a single rigid object. It may also be possible to incorporate optical flow information into the rigid object segmentation component. As before, feedback from higher levels, particularly the object tracking component of Level VI, may aid in the spatial correspondence and rigid object segmentation problems.

Level VI comprises motion and structure estimation or three-dimensional object tracking and is the major focus of this report. Once at least three noncollinear feature points, for example, are available for a single rigid object from Level V, three dimensional object state estimation may be initiated in order to further refine representations of the observed process. For the purposes of this investigation, solutions of the important problems of Levels I through V are assumed; the cameras are assumed to be stationary (although methods considered herein may immediately be extended

to the case of moving sensors), noisy feature position measurements as well as temporal and spatial correspondence information are assumed to be available, and only a single rigid object is present in the multiple-camera image sequences. These assumptions are very common in simulation studies as outlined in the literature review of Section 1.1.

Two additional levels in Figure 1.1 include knowledge refinement in Level VII and environmental interaction in Level VIII. Knowledge refinement refers to the accumulation and refinement of information from all previous levels in order to construct a meaningful representation of the observed processes. This decision-making level might be responsible, for example, for filter decoupling once structure estimates of rigid objects reach a pre-defined confidence level, fusion of rigidly attached objects, maintaining an estimate of relative motion or interaction of multiple objects, or perhaps object recognition and/or classification. The primary objective of most vision systems ultimately is to interact in some way with the environment or at least to form intelligent decisions concerning all observed processes. This is the function assigned to Level VIII and may include, for example, direct control of dynamic systems, path planning and navigation, threat detection with assessment and evasive strategy planning (collision avoidance), or perhaps compilation of surveillance information in security systems. Knowledge based systems will no doubt play a major role in forming solutions to these higher-level problems.

The hierarchical structure proposed in Figure 1.1 provides a means to proceed to higher levels once sufficient information is available, or revert to lower levels if insufficient information is present to proceed at the given level, particularly at Level VI. Generally, the proposed system is composed of three coupled multi-target tracking systems with increasing levels of complexity: in Level IV multiple features are tracked in the image planes; Level V fuses information from Level IV in order to track multiple features in three-dimensions; and Level VI is concerned with tracking multiple objects, composed of multiple features from Level V, in three dimensions with six degrees-of-freedom. It might be expected, therefore, that statistical-based multi-target tracking methods, which have been extensively studied in radar-based tracking applications, will be extremely useful in forming solutions to the vision-based motion analysis problem.

Another problem which has not been mentioned above, but is of fundamental importance in feature-based techniques, is that of feature occlusion due to object motion or other objects in the foreground of a scene. A common approach in simulation studies is to assume that occlusion does not occur [7]. However, recursive approaches, such as the Kalman filter, may provide natural solutions for state estimation in the presence of occlusion since measurements need not be of the same set of features at each observation event. Feature occlusion is an important component of the temporal and spatial correspondence problems and will not be explicitly examined in this report. In previous work [58] which is reviewed in Chapter 2, feature occlusion was treated by simply proceeding with state estimation, but in the absence of data from

occluded points.

Evaluation of methods for Level VI on real imagery is beyond the scope of current work. Such an analysis will require an intense review of feature detection and correspondence problems, sensor modeling and measurement noise characterization, and complex experimental arrangements in order to obtain long sequences (100 frames, for example) of digital, time-aligned, recorded imagery and ground truth data associated with object and/or camera motion. More complete implementations and performance analysis of proposed methods with real imagery constitute major proposals for future research outlined in Chapter 5. In this investigation, suitable imagery with or without occlusion is computer-generated for a translating and rotating block. Features of interest are the block corners whose positions in each image are perturbed by an additive pseudo-random noise process and made available to structure and motion estimation algorithms along with correspondence information. This approach to simulation analysis is consistent with similar studies in the literature [15] [28]-[34].

1.3 Chapter Summary

This chapter introduced and compared image flow and feature based approaches to motion analysis, and presented a survey of major contributions which apply extended Kalman filtering to motion and structure estimation. A possible hierarchical approach to feature-based motion analysis has been proposed in which motion and structure estimation represents only a single component. Major assumptions employed in this investigation that serve to isolate the motion and structure estimation problem have been discussed. The remainder of this report is organized as follows:

Chapter 2 treats an important but simplified problem in which the object of interest is constrained to lie on a known planar surface;

Chapter 3 treats the more general object tracking problem by introducing geometry and notation, developing state space models for three parameterizations of rotational motion, and presenting a simple single-frame filter initialization scheme;

Chapter 4 compares Cramer Rao bounds of the three parameterizations for a simplified trajectory, and presents simulation results and comparisons of estimation performance of the three filters for manoeuvring object trajectories;

Chapter 5 provides general discussions, conclusions, and proposals for future work.

Several appendices are provided so that this report may be self contained:

Appendix A provides a tutorial on parameterizations of relative orientation including the roll-pitch-yaw, Euler angle-axis, and quaternion representations, and equations which govern the temporal behavior of parameters;

Appendix B gives a review of Kalman and extended Kalman filtering for continuous- and discrete-time systems, as well as modifications such as local iterations;

Appendix C derives closed-form as well as recursive expressions representing the Cramer Rao lower bounds for discrete-time nonlinear state estimation;

Appendix D lists equations of measurement model Jacobians, which are required in the extended Kalman filter implementation, for the three parameterizations of rotational motion.

2. Motion On A Known Planar Surface

In a large number of motion analysis problems, particularly in many robotics applications, objects of interest are constrained to lie on a surface which is known or can be approximated analytically. This chapter briefly introduces the Kalman filtering approach by reviewing previous research by the author, [58, 59]. This work proposed linear motion and structural models for a rigid object which is constrained to lie on a known planar surface, and a nonlinear measurement model which describes noisy monocular (single camera) perspective projection observations from an arbitrary position above the known plane. As in the work of Broida *et al.* [32], object motion was assumed to be of constant translational and rotational velocity. These assumptions, although restrictive, provided a simplified framework in which to investigate the performance of Kalman filtering methods in motion analysis problems.

This chapter first introduces notation and geometry for reference frames, coordinate transformations, and the perspective projection image formation model in Section 2.1. Section 2.2 briefly reviews the planar motion method by developing structure, motion, and measurement models, and presenting a reduced set of simulation results. Section 2.3 summarizes and discusses the methods and results of this chapter.

2.1 Notation, Geometry and Perspective Projection

Planar motion and object structure are recovered from positions of object feature points extracted from a monocular image sequence. Three reference frames, denoted

$$\begin{aligned} F_O &\sim \text{Object-fixed,} \\ F_E &\sim \text{Earth-fixed, and} \\ F_C &\sim \text{Camera-fixed} \end{aligned} \tag{2.1}$$

are used in this problem.

A reference frame, F_α, where $\alpha \in \{O, E, C\}$, is defined by a point O_α, the origin

of F_α, and three orthonormal vectors $\mathbf{i}_\alpha$, $\mathbf{j}_\alpha$, and $\mathbf{k}_\alpha$ which lie along the x_α, y_α, and z_α axis of F_α, respectively. A vector $\mathbf{r}$ which describes the position of a feature point on the object with respect to O_α and is expressed in F_α is denoted by $\mathbf{r}_\alpha$. The components of $\mathbf{r}_\alpha$ are written as $\mathbf{r}_\alpha = [r_{\alpha,1}, r_{\alpha,2}, r_{\alpha,3}]^T_\alpha$.

A second reference frame, F_β, is related to F_α by a translation $\mathbf{T}^\alpha_\beta$, which represents a vector from O_α to O_β and, when expressed with respect to F_α, is written as $[\mathbf{T}^\alpha_\beta]_\alpha$, and an orthogonal 3×3 matrix, $\mathbf{I}^\alpha_\beta$, which denotes a transformation from the α-basis to the β-basis and represents a rigid rotation of $\mathbb{R}^3$. Appendix A provides a tutorial review of the Euler angle-axis, quaternion, and roll-pitch-yaw parameterizations for basis transformations represented by $\mathbf{I}^\alpha_\beta$. The planar method employs only the yaw angle of the roll-pitch-yaw parameterization. With this notation,

$$[\mathbf{r}_\alpha]_\beta = \mathbf{I}^\alpha_\beta \mathbf{r}_\alpha, \text{ and} \tag{2.2}$$

$$\mathbf{r}_\beta = [\mathbf{r}_\alpha]_\beta + \left[\mathbf{T}^\beta_\alpha\right]_\beta. \tag{2.3}$$

If optical distortions of an imaging device are not severe, a perspective projection image formation model can be used to relate the observed location of a feature in a two-dimensional image to the three-dimensional coordinates of that feature in a camera-centred reference frame. Figure 2.1 shows the assumed geometry for the image formation model. The position of the ith feature point, p^i, with respect to a camera-centred reference frame, F_C, is described by the vector $\mathbf{r}^i_C$. A feature at point p^i is projected onto the image plane and appears at point (L, u^i, v^i) in F_C, where L is the effective focal length. The origin of F_C is taken at the centre of projection of the imaging lens. The $\mathbf{x}_C$ axis of F_C is aligned with the optical axis, while $\mathbf{y}_C$ is parallel to a scan line and positive to the right in the image. The $\mathbf{z}_C$ axis is positive down in the image to form a right-handed coordinate system. Neglecting separate horizontal and vertical scale factors, and assuming any distortion effects are negligible, the projected image plane position of feature point p^i is given by

$$u^i = L\left(\frac{r^i_{C,2}}{r^i_{C,1}}\right), \text{ and}$$

$$v^i = L\left(\frac{r^i_{C,3}}{r^i_{C,1}}\right). \tag{2.4}$$

The central projection transformation defined in (2.4) demonstrates the loss of depth information in monocular imaging systems; the image position (u^i, v^i) is invariant under any scalar multiple of $\mathbf{r}^i_C$.

Figure 2.2 shows a typical physical arrangement for the planar motion problem. The object of interest, shown here as a block, is described by feature points, for example the block corners, in F_O. The earth-fixed reference frame, F_E, is defined such that the x_E-y_E coordinate plane is the plane of motion. The x_O-y_O plane of F_O is defined to be parallel to the plane of motion in which case the axis of rotation

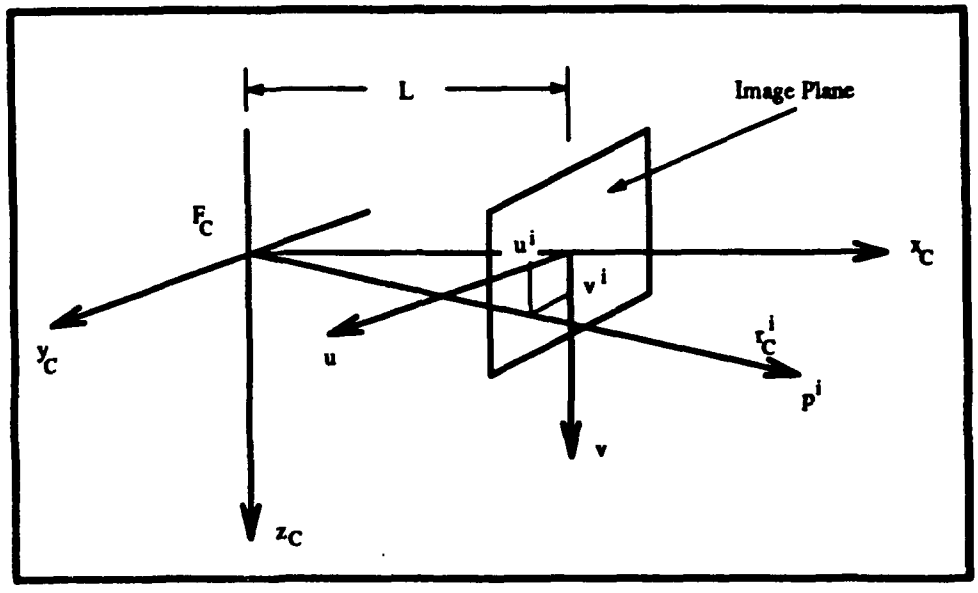

Figure 2.1

Central projection image formation geometry. A feature at point p^i projects onto the point (L, u^i, v^i) in the camera-centred frame, F_C, where L is the effective focal length.

is parallel to $\mathbf{z}_E$. The position and orientation of F_C is assumed to be known with respect to F_E.

2.2 Estimation of Structure and Planar Motion

The extended Kalman filtering (EKF) approach requires first that dynamical and measurement models be expressed in state space form. Appendix B provides a review of the filtering equations which, in general, are based on a dynamical (plant or process) model of the form

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t)] + \mathbf{G}(t)\mathbf{w}(t), \tag{2.5}$$

where the white noise process[1] $\mathbf{w}(t) \sim N(\mathbf{0}, \mathbf{Q}(t))$ is mapped to the state space through the matrix G. The measurement model is expressed in discrete-time form, since images arrive at discrete instants in time, as

$$\mathbf{z}(k) = \mathbf{h}[\mathbf{x}(t_k)] + \mathbf{v}(k), \tag{2.6}$$

where the measurement noise process, $\mathbf{v}(k) \sim N(\mathbf{0}, \mathbf{R}(k))$, is also assumed to be temporally white. Motion, structure, and measurement models for the planar motion problem are required in the above form.

[1]The notation N(m, C) denotes a jointly Gaussian population with mean vector m and covariance matrix C.

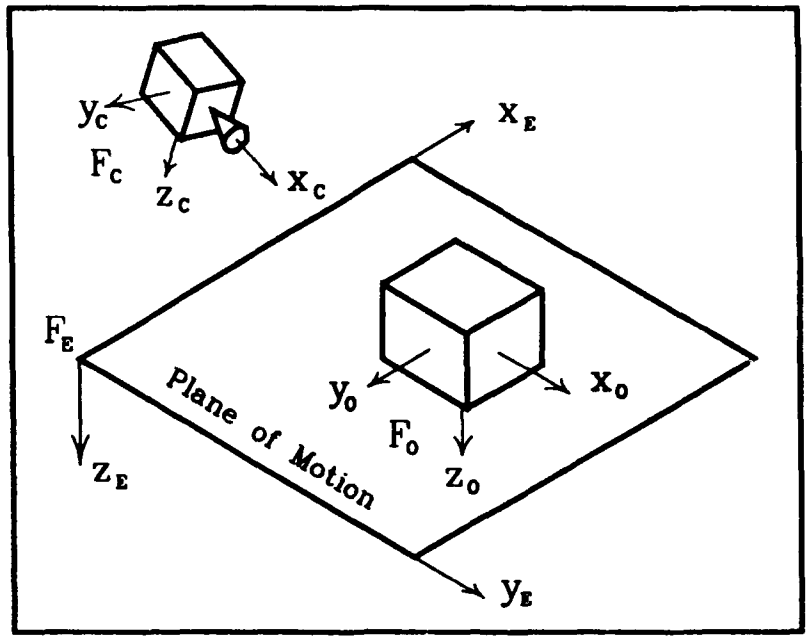

Figure 2.2

System geometry for planar motion analysis.

Motion and Structure Models: The position of the origin of F_O relative to F_E is represented by $[T_O^E]_E(t) = [x(t), y(t), 0]_E^T$. The orientation of F_O relative to F_E is described by the yaw angle $\psi(t)$ whose (constant) time derivative w defines the rotational velocity vector as $\mathbf{w} = [0, 0, w]_E^T$. The state vector for object motion is then given by

$$\mathbf{x}_m(t) \triangleq [x(t), \dot{x}, y(t), \dot{y}, \psi(t), w]^T , \tag{2.7}$$

where superscript 'T' denotes the matrix transpose.

The origin of F_O is defined to have constant translational velocity and lies at the intersection of the rotational axis and the plane of motion. Note in this case that the object-centred frame depends on the assumed nature of object motion, which results in unobservability in the absence of rotational motion. Further comments on this problem are given in Section 3.2. It is also assumed that one feature point is observed whose height above the plane of motion is known or estimated. This point is fixed in the (vertical) $\mathbf{x}_O$-$\mathbf{z}_O$ plane. Hence both the $\mathbf{z}_O$- and $\mathbf{y}_O$-coordinates of this special feature point are known and removed from the estimation process. In order to simplify notation in the following development, removal of two coordinates of one feature point is neglected.

The full state vector which describes object motion and structure is defined as

$$\mathbf{x}(t) = \begin{bmatrix} \mathbf{x}_m(t) \\ \mathbf{x}_s \end{bmatrix}, \tag{2.8}$$

where $\mathbf{x}_s$ is the structure state vector,

$$\mathbf{x}_s \triangleq \left[(\mathbf{r}_O^1)^T, \ldots, (\mathbf{r}_O^{N_f})^T\right]^T, \tag{2.9}$$

and N_f is the number of feature points of interest. A linear discrete-time difference equation which describes time propagation of the state $\mathbf{x}(t)$ over a sample period T_k can be written as

$$\mathbf{x}(k+1) = \mathbf{\Phi}(k)\mathbf{x}(k) + \mathbf{G}(k)\mathbf{w}(k), \tag{2.10}$$

where

$$\mathbf{\Phi}(k) \triangleq \begin{bmatrix} \mathbf{\Phi}_m(k) & \mathbf{0}_{6\times 3N_f} \\ \mathbf{0}_{3N_f\times 6} & \mathbf{I}_{3N_f} \end{bmatrix}, \tag{2.11}$$

in which $\mathbf{0}_{p,q}$ represents a $p \times q$ block of zeros, $\mathbf{I}_p$ is the $p \times p$ identity matrix, and

$$\mathbf{\Phi}_m(k) \triangleq \begin{bmatrix} 1 & T_k & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & T_k & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & T_k \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.12}$$

Higher-order motion may be incorporated into the system model which remains linear, due to the planar motion assumption, in the elements of the state vector. The overall state transition matrix $\mathbf{\Phi}(k)$ in (2.11) is extremely sparse which leads to very efficient time propagation in the discrete-time extended Kalman filter implementation.

Measurement Model: The instantaneous position and orientation of F_O relative to F_E are given by the translation

$$\left[\mathbf{T}_O^E\right]_E(t) = [x(t), y(t), 0]_E^T, \tag{2.13}$$

and basis transformation matrix

$$\mathbf{I}_E^O(\psi) = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{2.14}$$

while the known position and orientation of F_C relative to F_E are defined by $\mathbf{T}_C^E$ and $\mathbf{I}_C^E$, respectively. The position of the ith feature point with respect to F_C is written as

$$\mathbf{r}_C^i = \mathbf{I}_C^E\left\{\mathbf{I}_E^O\mathbf{r}_O^i + \left[\mathbf{T}_O^E\right]_E - \left[\mathbf{T}_C^E\right]_E\right\}. \tag{2.15}$$

which can be expressed in component form (see Appendix A) as

$$\begin{aligned} r^i_{C,1} &= [\mathbf{i}_C]^T_E \left\{ \mathbf{I}^O_E \mathbf{r}^i_O + \left[\mathbf{T}^E_O\right]_E - \left[\mathbf{T}^E_C\right]_E \right\}, \\ r^i_{C,2} &= [\mathbf{j}_C]^T_E \left\{ \mathbf{I}^O_E \mathbf{r}^i_O + \left[\mathbf{T}^E_O\right]_E - \left[\mathbf{T}^E_C\right]_E \right\}, \\ r^i_{C,3} &= [\mathbf{k}_C]^T_E \left\{ \mathbf{I}^O_E \mathbf{r}^i_O + \left[\mathbf{T}^E_O\right]_E - \left[\mathbf{T}^E_C\right]_E \right\}. \end{aligned} \tag{2.16}$$

With N_f feature points of interest the measurement model is given by

$$\mathbf{z}(k) = \mathbf{h}[\mathbf{x}(k)] + \mathbf{v}(k), \tag{2.17}$$

with

$$\mathbf{h}[\mathbf{x}(k)] = \begin{bmatrix} \mathbf{h}^1[\mathbf{x}(k)] \\ \mathbf{h}^2[\mathbf{x}(k)] \\ \vdots \\ \mathbf{h}^{N_f}[\mathbf{x}(k)] \end{bmatrix}, \tag{2.18}$$

where the image plane position, $\mathbf{h}^i[\mathbf{x}(k)]$, of the ith feature point is computed with the perspective projection transformation of (2.4),

$$\mathbf{h}^i[\mathbf{x}(k)] = \begin{bmatrix} h^i_1[\mathbf{x}(k)] \\ h^i_2[\mathbf{x}(k)] \end{bmatrix} = L \begin{bmatrix} \frac{r^i_{C,2}}{r^i_{C,1}} \\ \frac{r^i_{C,3}}{r^i_{C,1}} \end{bmatrix}, \tag{2.19}$$

and $\mathbf{v}(k)$ models the measurement noise which is assumed to a discrete-time, zero-mean, white Gaussian process with covariance matrix $\mathbf{R}(k)$.

A single image provides $2N_f$ equations in $6+3N_f-2$ unknowns (since one feature point has two known coordinates). Each additional feature point provides two more equations, but increases the dimension of the state vector by three unknowns. With four feature points, for example, each image provides 8 equations in 16 unknown variables. In this case, at least two image frames are required to obtain more equations than unknowns. Three of these variables (x, y, ψ) are time varying while all others are constant.

Simulation Results: This discrete-time state estimation problem is formulated with a linear dynamic model and a nonlinear measurement model. The discrete-time extended Kalman filter (EKF) was found to yield satisfactory results without local iterations when object structure is known, but demonstrates somewhat slower convergence in the absence of prior structural information. The iterated, extended Kalman filter (IEKF), which is also reviewed in Appendix B, generally gives significant improvements in the rate of convergence while structure is being estimated. In extensive simulations [58], the IEKF demonstrated reliable convergence to true parameters with a simple single-frame initialization of object position and orientation, and with all other parameters set to zero. This Section, however, shows Monte Carlo

simulation results based on 60 runs with random filter initialization taken from a Gaussian population centred about the true initial state, i.e., simulations assume the availability of an unbiased estimator of $\mathbf{x}(0)$ with statistics $\mathbf{x}_0 \sim \mathrm{N}(\mathbf{x}(0), \mathbf{P}_0)$, where the covariance matrix $\mathbf{P}_0$ defines the level of prior information in the initial estimate $\mathbf{x}_0$. Filter initialization is then taken as $\hat{\mathbf{x}}(0|-1) = \mathbf{x}_0$ and $\hat{\mathbf{P}}(0|-1) = \mathbf{P}_0$. Filter mean errors (ME) and root-mean-square-errors (RMSE) are computed immediately following each observation event over the 60 simulation runs.

The EKF is generally a nonlinear, suboptimal state estimation technique. A powerful result which provides a performance assessment of parameter estimation techniques is the Cramer-Rao inequality [79]. Appendix C reviews the significance and derivation of Cramer-Rao lower bounds (CRLB) for this estimation problem. Generally, the error covariance matrix $\mathbf{S}(N)$ of an *unbiased* estimator $\hat{\mathbf{x}}(N)$ of the state $\mathbf{x}(N)$, where the estimate is based on the prior information in $\mathbf{x}_0$ and the observations $\mathbf{z}(0), \ldots, \mathbf{z}(N)$, is lower-bounded by the inverse of Fisher's information matrix [80, pp. 91-93]. This result, in theory, provides estimation error variance lower bounds for elements of $\hat{\mathbf{x}}(N)$. It should be noted, however, that these bounds assume an unbiased estimation procedure; a state dependent bias can result in higher or lower bounds [78] (see Appendix C). As a result, most analysis [28]-[31], [61], [62], proceed under the assumption of unbiased estimation resulting in approximate covariance bounds.

The geometry of Figure 2.2 was used to generate sequences of image feature points with occlusion and correspondence information. The camera position and orientation is defined by a translation $[\mathbf{T}_C^E]_E = [0, -30, -30]^T$, and the transformation $\mathbf{I}_C^E$ formed from Euler angles $\psi_C = 45°$ (yaw or pan), $\theta_C = -45°$ (pitch or tilt), and $\phi_C = 0$ (roll about optical axis).

The object is described by eight corners of a solid rectangular block 20 units wide and 10 units in height and depth. Four feature points labelled p^1, p^3, p^5, and p^7 define the structure of the block. It was found that Cramer-Rao bounds do not decrease significantly when more than four feature points are considered. Points p^1 and p^5 are diagonally opposite corners on the bottom face of the block, and hence are subject to occlusion due to object motion. Points p^3 and p^7 are diagonally opposite corners on the top face of the block and are diagonally opposite to p^1 or p^5 on side faces. Point p^5 (sometimes occluded) is selected as the special feature point which is known to lie in the plane of motion and is further assumed to lie on the negative $\mathbf{x}_O$ axis. All structure parameters are retained in the state vector during occlusion.

The true trajectory is defined by an initial position $(x(0), y(0)) = (30, -25)$ units, orientation $\psi(0) = 2.1$ radians, constant translational velocity $(\dot{x}, \dot{y}) = (2.5, 5.0)$ units/s, and constant angular velocity $w = -0.5$ radians/s. Initial estimate uncertainty is defined by a diagonal $\mathbf{P}_0$ in which the square roots of the diagonal entries represent standard deviations of 5 units for position, 3 units/s for translational velocity, 0.35 radians for angular position, 1.0 radians/s for angular velocity, and 5 units

for each of the 10 structural parameters (the y- and z-coordinates of p^5 are not estimated). Local iterations at each observation event are employed with a maximum of 10 and terminating if the maximum element of the difference in successive iterates is less than 0.01. In most cases, only 5 to 7 iterations were required during initial observation events, but as the filter accumulates information, only one or two iterations were necessary.

Figure 2.3 shows a true and a sample noisy image sequence of feature point trajectories with additive measurement noise of standard deviation $\sigma_v = 0.02$ units and unity focal length. A 20 second time interval is employed with a sample period of $T = 0.1$ seconds. The measurement noise level of 0.02 units represents quantities on the order of 5% to 10% of the object image size when the object is closest to the camera, but as the object moves away from the camera noise levels are on the order of 11% to 23% of the object image size at the end of the trajectory. A feature point occlusion map is shown in Figure 2.4. Feature point p^1 is visible in the first 9 image samples or almost 1 second before being occluded for about 3 seconds. The special feature point p^5 is occluded during time periods from 6.5 to 10 seconds and from about 19 seconds to the end of the trajectory.

Figures 2.5 and 2.6 show the root-mean-squared error (RMSE), Cramer-Rao lower bound (CRLB), and mean error (ME) results for a 60 run Monte Carlo simulation. Figure 2.5 shows results for motion parameter estimation. Position and orientation estimates tend to be biased over the first 10 seconds as indicated by the ME trace. After 10 seconds, the RMSE for x-position approaches and follows the CRLB. RMS errors for y-position and velocity are also initially biased, but approach and follow the CRLB just after 4 seconds. Angular position RMSE remains above the CRLB throughout the trajectory and tends to increase slightly during the occlusion of p^5, but demonstrates only small bias after 10 seconds. Good performance is shown in the estimation of angular velocity and x-velocity.

Figure 2.6 shows results for structure estimation of one point, p^3, on the top of the block (not subject to occlusion), and both points, p^1 and p^5, on the bottom of the block. Satisfactory RMSE performance is demonstrated for p^3 structure after 5 seconds with negligible bias. Simulation results for estimation of p^7 structure (not shown) are almost identical to those shown for p^3. Errors in estimates of p^1 structure during occlusion may be largely responsible for bias in other state estimates over the initial portion of the trajectory. The presence of p^1 in images sampled after $t = 4$ seconds results in a rapid decrease in structure RMS errors as estimation performance approaches the CRLB. Estimation of the x_O-coordinate of p^5 is also somewhat biased particularly just after p^1 becomes visible and during the time interval $6 < t < 10$ seconds when p^5 itself is occluded. RMS errors of all structural parameters approach and follow the CRLB towards the end of the trajectory.

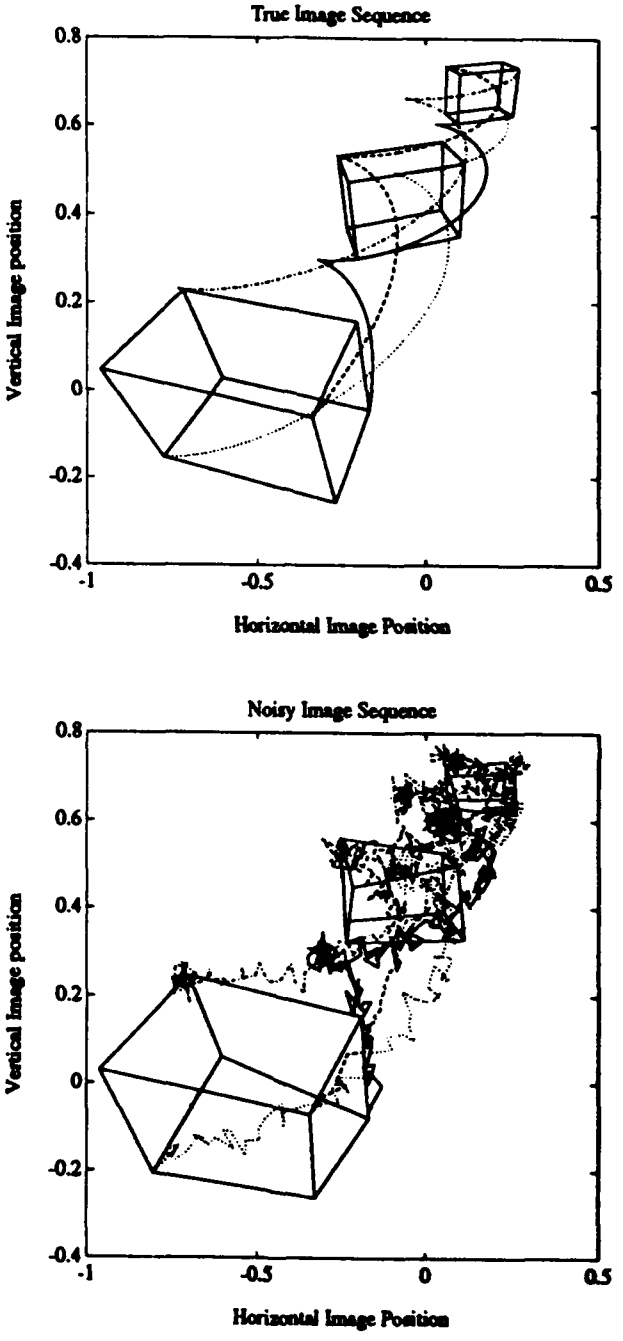

Figure 2.3

True and sample noisy image feature point trajectories for planar motion.

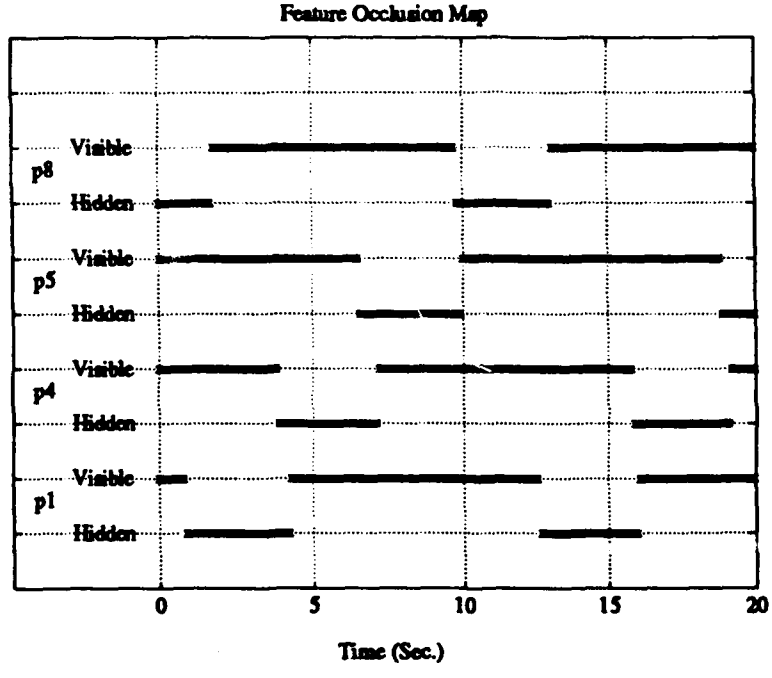

Figure 2.4

Feature point occlusion map.

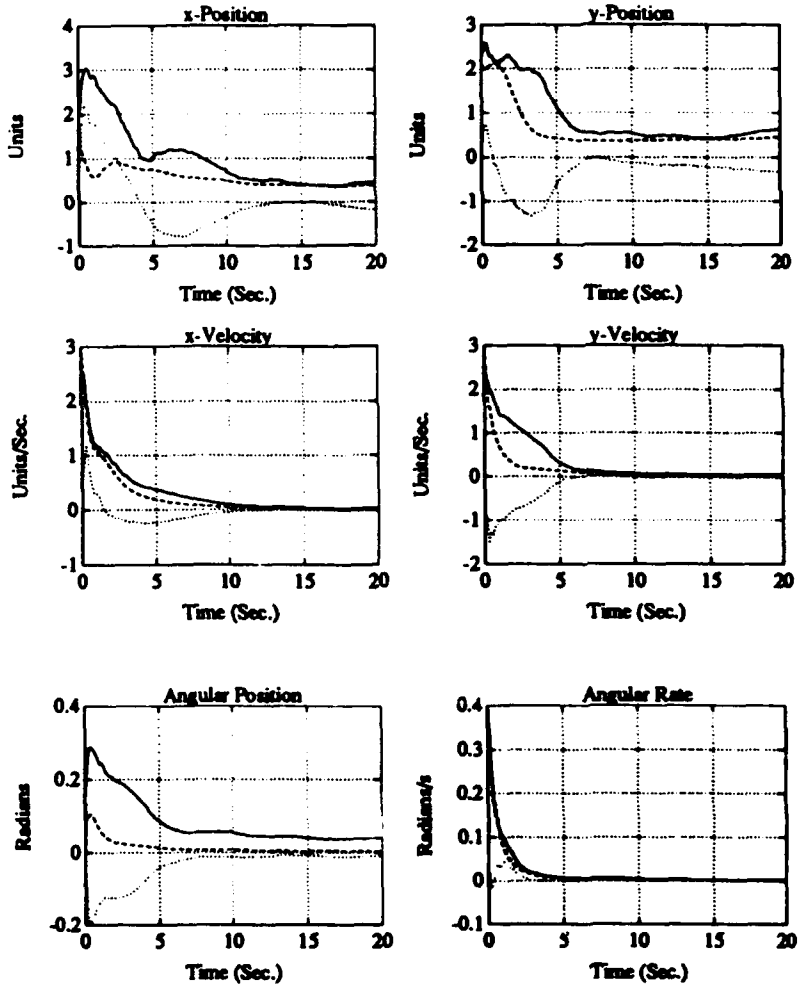

Figure 2.5

Planar motion parameter Monte Carlo simulation results. All plots show —— RMSE, – – – CRLB, and ····· ME.

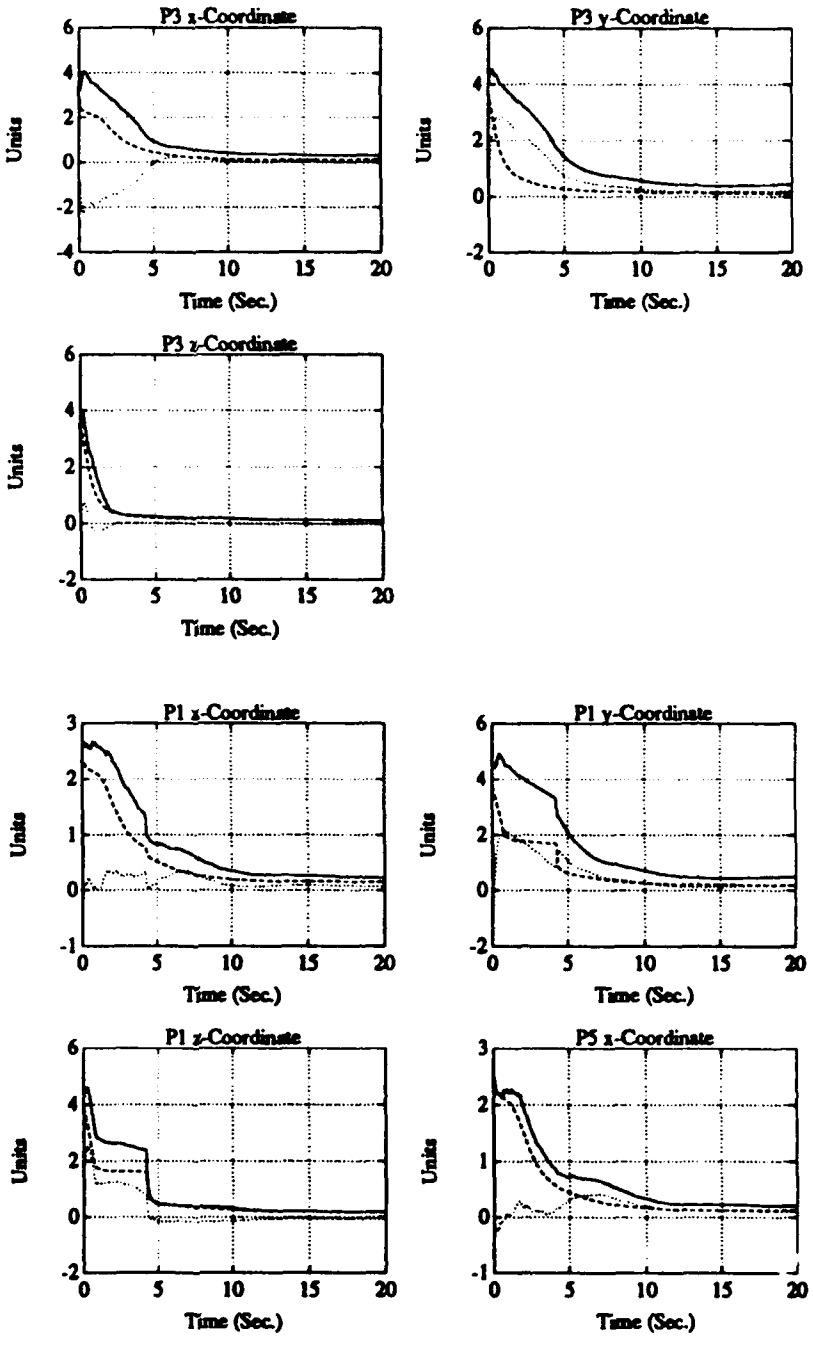

Figure 2.6

Planar motion, structure parameter Monte Carlo simulation results.
All plots show —— RMSE, – – – CRLB, and ····· ME.

2.3 Chapter Summary

This chapter has introduced notation for reference frames and coordinate transformations that will be used extensively in following chapters and in the Appendices. The perspective projection image formation model is fairly standard in image analysis work, with the exception of some methods which are based on orthographic projection—a simplification which leads to mathematically tractable systems of equations in some cases. Only perspective projections will be considered in this work.

Previous research proposed motion, structure and measurement models which describe motion of a rigid object constrained to a known planar surface with monocular observations from above the known plane. Monte Carlo simulation results compared to approximate Cramer-Rao bounds demonstrated good performance in recovery of absolute motion and structure through iterated extended Kalman filtering which parallels the approach of Broida *et al.* [32]. This investigation demonstrated filter performance that might be expected for constant translational and rotational velocity. A fundamental assumption of this technique is that the height of one feature point above plane is known or approximated. Another important limitation of the proposed technique and methods of Broida *et al.* [32] and Young *et al.* [35] is that the systems as formulated are not observable for zero angular velocity. The absence of rotational motion must be detected during initialization, and an alternate parameterization is required in which one feature point is selected as the origin of the object-fixed frame and angular velocity is removed from the state vector. Filter performance, in fact, was found to be seriously degraded as rotational rates become very small.

Current research, which is the subject of subsequent chapters of this report, removes assumptions regarding partial knowledge of the position of some special feature point, extends motion models to six-degree-of-freedom manoeuvring object trajectories, extends the measurement model to multiple imaging systems, and investigates three parameterizations of rotational motion. In addition, the object-centred frame is defined independent of object motion so that the system is observable, provided at least three feature points are available, for arbitrary object motion.

3. Approximate Models for General Motion

Methods for planar motion considered in the previous chapter employed several assumptions which may be difficult to satisfy or verify in analysis of general scenes. This chapter develops approximate dynamic models for six-degree-of-freedom manoeuvring object motion, and a perspective projection measurement model for multiple-camera imaging systems. As in the planar motion problem, dynamical and measurement models must be obtained in state space form as

$$\dot{\mathbf{x}}(t) = \mathbf{f}[\mathbf{x}(t)] + \mathbf{G}(t)\mathbf{w}(t), \tag{3.1}$$

$$\mathbf{z}(k) = \mathbf{h}[\mathbf{x}(t_k)] + \mathbf{v}(k), \tag{3.2}$$

where the process noise $\mathbf{w}(t) \sim N(\mathbf{0}, \mathbf{Q}(t))$, which is mapped to the state space through the matrix $\mathbf{G}$, and the measurement noise, $\mathbf{v}(k) \sim N(\mathbf{0}, \mathbf{R}(k))$, are assumed to be temporally white.

Notation and assumed geometry are introduced in Section 3.1. Section 3.2 discusses the structural model for a rigid object. The measurement model for observations with multiple imaging systems is presented in Section 3.3. Translational motion models are developed in Section 3.4, and approximate dynamic models for the Euler angle-axis, roll-pitch-yaw, and quaternion parameterizations of rotational motion are proposed in Section 3.5. A simple initialization scheme based on a single measurement event is the subject of Section 3.6. A brief chapter summary is provided in Section 3.7.

3.1 Notation and Geometry

This chapter employs the same notation for reference frames and coordinate transformations as presented in Section 2.1, but with multiple-camera image sequences,

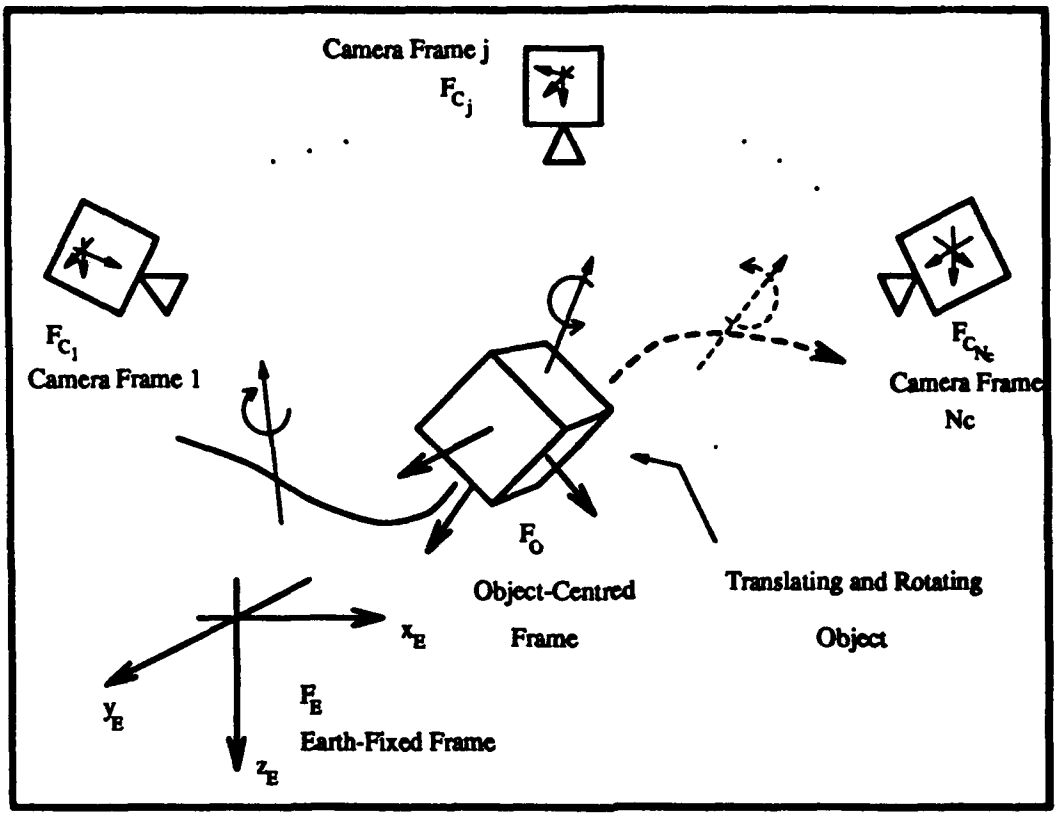

Figure 3.1

System geometry for general motion analysis.

multiple camera reference frames are required. Reference frames denoted as

$$\begin{aligned} F_O &\sim \text{Object-fixed,} \\ F_E &\sim \text{Earth-fixed, and} \\ F_{C_j} &\sim j\text{th Camera-fixed, } j = 1,2,\ldots,N_C \\ &\quad \text{for } N_C\text{-ocular imaging systems,} \end{aligned} \tag{3.3}$$

are used in the following development.

Figure 3.1 illustrates a typical physical arrangement for the problem considered in this work. Two or more cameras observe a single rigid object whose motion is both translational and rotational. Image formation and axis orientation for each camera frame is as described in Figure 2.1 of Section 2.1. The position, $[\mathbf{T}^E_{C_j}]_E$, and orientation, $\mathbf{I}^E_{C_j}$, of the jth camera frame, F_{C_j}, $j = 1,2,\ldots,N_C$, is known or estimated with respect to the stationary earth fixed frame, F_E. Object structure is again formulated in terms of positions, $\mathbf{r}^i_O$, of feature points expressed with respect to the object-centred frame, F_O, while object motion, which consists of translational motion of the origin of F_O together with rotational motion of F_O, is described with respect to the stationary earth-fixed frame F_E. As a result, parameterization of

object motion and structure is independent of position, orientation, and motion of the imaging systems.

The overall state vector $\mathbf{x} \in \mathbb{R}^N$ includes states corresponding to translational motion, rotational motion and object structure:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_t \\ \mathbf{x}_r \\ \mathbf{x}_s \end{bmatrix}, \tag{3.4}$$

where

$$\begin{aligned} \mathbf{x}_t \in \mathbb{R}^{N_t} &\sim \text{Translational motion state vector,} \\ \mathbf{x}_r \in \mathbb{R}^{N_r} &\sim \text{Rotational motion state vector,} \\ \mathbf{x}_s \in \mathbb{R}^{N_s} &\sim \text{Structure state vector.} \end{aligned} \tag{3.5}$$

The dynamical model is written in the general form

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \mathbf{f}_t[\mathbf{x}_t(t)] \\ \mathbf{f}_r[\mathbf{x}_r(t)] \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{G}_t & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_r \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{w}_t \\ \mathbf{w}_r \end{bmatrix}, \tag{3.6}$$

where subscripts 't', and 'r' denote translational and rotational motion models, respectively. As indicated, elements of the structure state vector are constant in time.

3.2 Structure Model

The rigid object is defined by N_f feature points, p^i, $i = 1, 2, \ldots, N_f$, whose position vectors with respect to F_O, $\mathbf{r}_O^i$, collectively define the structure set $\mathcal{S} = \{\mathbf{r}_O^1, \mathbf{r}_O^2, \ldots, \mathbf{r}_O^{N_f}\}$. Observation of at least three noncollinear feature points by at least two cameras at each observation event provides sufficient information to fix an object-centred frame on the object. A number of techniques, including [28]-[38], [58, 59], define an object-centred frame whose orientation and position with respect to the object itself is based in part on the assumed motion of the object. For example, in Chapter 2, the origin of the object-centred frame is defined to be a point which satisfies the following constraints:

1. Rigid object assumption - The position of the origin with respect to all observed feature points is constant in time;

2. Planar motion assumption - The origin lies in the known plane of motion; and

3. Constant velocity assumption - The origin moves with constant translational velocity (on the axis of rotation) with respect to the earth-fixed frame.

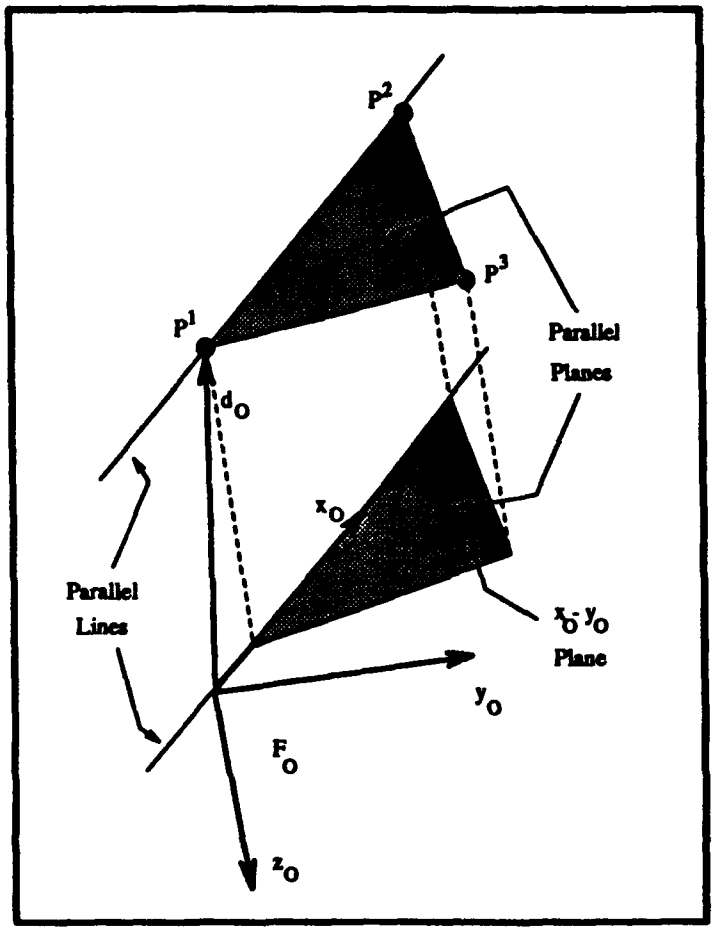

Figure 3.2

Fixing the object-centred frame with three feature points.

This leads to unobservability even if the assumption of constant translational velocity is correct in the absence of rotational motion since in this case an infinite number of points will satisfy items 1-3 above. Similar arguments can also be applied to methods proposed in [28]-[38] which cannot be directly employed in the case of manoeuvring objects for which the true nature of translational and rotational motion is unknown. In this case, the object-centred frame must be independent of object motion and defined on the basis of observed feature points only.

In the present work, F_O is fixed on the object by selecting, as will be discussed further in Section 3.6, three noncollinear feature points $\{p^1, p^2, p^3\}$, (three noncollinear image points are necessarily projections of three noncollinear 3D points) and imposing the following constraints as demonstrated in Figure 3.2:

1. Feature point p^1 lies at a known position $\mathbf{d}_O = [d_{O,1}, d_{O,2}, d_{O,3}]^T$ in F_O, and hence its structure vector is known;

2. One of the coordinate axis of F_O, say the $\mathbf{x}_O$ axis, is parallel to the line passing through p^1 and p^2, in which case the y_O- and z_O-coordinates of the structure vector of p^2 are known to be $d_{O,2}$ and $d_{O,3}$, respectively; and

3. One of the coordinate planes of F_O, say the $\mathbf{x}_O$-$\mathbf{y}_O$ plane, is parallel to the plane containing p^1, p^2, and p^3, in which case the z_O-coordinate of the structure vector of the third point is known to be $d_{O,3}$.

Selection of the vector $\mathbf{d}_O$, which is fixed at the time of filter initialization and thereafter remains constant, will also be discussed in Section 3.6. This method results in one of four possible orientations of F_O with respect to the three points: feature point p^2 in item 2 could have a positive or negative $\mathbf{x}_O$ coordinate, and/or feature point p^3 in item 3 could have a positive or negative $\mathbf{y}_O$ coordinate. These four discrete solutions are related by diagonal basis transformations (rotations of the basis set by 0 or π about a coordinate axis). The three special feature points would most likely be selected, as outlined in Section 3.6, from a larger set of all available feature points, which leads to many possible positions and orientations of F_O with respect to the object. Once the initial ordering of three special points has been specified, depending on the initialization error and filter transients, any one of the four possible orientations of F_O is possible. Obviously, the estimation process is likely to be poorly behaved if images of these three points are very close together relative to measurement noise levels. The important advantage, however, is that F_O is defined independent of object motion.

The approach taken here is to use measurements of feature positions in the images directly, and enforce the three constraints above by fixing the six known structural coordinates indicated in items 1,2, and 3 by excluding these parameters from the estimation process. Explicit removal of the six structural coordinates from the structure state vector in the development of dynamic and measurement models below would lead to unnecessarily complicated notation. For brevity, the full structure state vector is indicated by the notation, but it is understood that these six coordinates are removed from the estimation process. The structure state vector, $\mathbf{x}_s \in \mathbb{R}^{N_s}$, where $N_s = 3N_f$ is written as

$$\mathbf{x}_s = \begin{bmatrix} \mathbf{r}_O^1 \\ \mathbf{r}_O^2 \\ \vdots \\ \mathbf{r}_O^{N_f} \end{bmatrix}. \tag{3.7}$$

For a rigid object, the elements of $\mathbf{x}_s$ are constant in time.

At any given sample time, only a subset of feature points $\mathcal{S}_{v_j} \subseteq \mathcal{S}$ will be visible in the image of the jth camera due to occlusion of feature points, i.e., the same set of feature points need not be observed by all cameras at each measurement event. In addition, the total number of feature points, N_f, is generally an integer-valued function of time; as new object features become of interest, they might be added

to the structure set, and when structure estimates reach some pre-defined threshold confidence level, they might be removed from the structure set and the estimation process. These important aspects of motion analysis are not explicitly examined in this report. In the present work, it is assumed that at least three noncollinear feature points are observed by at least two cameras at each observation event, and that the same three points are available without occlusion over the simulation time interval. These assumptions should be viewed as simplifications for development and presentation of the problem under consideration rather that restrictions imposed by the proposed methods.

3.3 Measurement Model

The measurement model is formed from a set of N_C individual measurement models which have identical form and each corresponds to a single camera. The instantaneous position and orientation of F_O relative to F_E are given by the translation

$$\left[\mathbf{T}_O^E\right]_E(t) \equiv \mathbf{p}(t) = [x(t), y(t), z(t)]_E^T, \tag{3.8}$$

and basis transformation $\mathbf{I}_E^O(\zeta)$, respectively, in which elements of the vector ζ depend on the parameterization of rotational motion. The known instantaneous position and orientation of the jth camera frame, F_{C_j}, with respect to F_E is specified by a translation vector $\mathbf{T}_{C_j}^E$ and basis transformation $\mathbf{I}_{C_j}^E$, respectively. The position of the ith feature point with respect to F_{C_j} can be written as

$$\mathbf{r}_{C_j}^i = \mathbf{I}_{C_j}^E \left\{\mathbf{I}_E^O \mathbf{r}_O^i + \left[\mathbf{T}_O^E\right]_E - \left[\mathbf{T}_{C_j}^E\right]_E\right\}, \tag{3.9}$$

and writing

$$\mathbf{I}_{C_j}^E = \left[[\mathbf{i}_{C_j}]_E, [\mathbf{j}_{C_j}]_E, [\mathbf{k}_{C_j}]_E\right]^T, \tag{3.10}$$

gives $\mathbf{r}_{C_j}^i$ in component form as

$$\begin{aligned} r_{C_j,1}^i &= \left[\mathbf{i}_{C_j}\right]_E^T \left\{\mathbf{I}_E^O \mathbf{r}_O^i + \left[\mathbf{T}_O^E\right]_E - \left[\mathbf{T}_{C_j}^E\right]_E\right\}, \\ r_{C_j,2}^i &= \left[\mathbf{j}_{C_j}\right]_E^T \left\{\mathbf{I}_E^O \mathbf{r}_O^i + \left[\mathbf{T}_O^E\right]_E - \left[\mathbf{T}_{C_j}^E\right]_E\right\}, \\ r_{C_j,3}^i &= \left[\mathbf{k}_{C_j}\right]_E^T \left\{\mathbf{I}_E^O \mathbf{r}_O^i + \left[\mathbf{T}_O^E\right]_E - \left[\mathbf{T}_{C_j}^E\right]_E\right\}. \end{aligned} \tag{3.11}$$

With N_j observed feature points, the measurement model for the jth camera is given by

$$\mathbf{z}_j(k) = \mathbf{h}_j[\mathbf{x}(k)] + \mathbf{v}_j(k), \tag{3.12}$$

where

$$\mathbf{h}_j[\mathbf{x}(k)] = \begin{bmatrix} \mathbf{h}_j^1[\mathbf{x}(k)] \\ \mathbf{h}_j^2[\mathbf{x}(k)] \\ \vdots \\ \mathbf{h}_j^{N_j}[\mathbf{x}(k)] \end{bmatrix}, \tag{3.13}$$

in which subscripts refer to the camera and superscripts refer to the feature point. The image plane position, $\mathbf{h}_j^i[\mathbf{x}(k)]$, of the ith feature point in the jth camera is computed through perspective projection as

$$\mathbf{h}_j^i[\mathbf{x}(k)] = \begin{bmatrix} h_{j,1}^i[\mathbf{x}(k)] \\ h_{j,2}^i[\mathbf{x}(k)] \end{bmatrix} = L_j \begin{bmatrix} \frac{\mathbf{r}_{C_j,2}^i}{\mathbf{r}_{C_j,1}^i} \\ \frac{\mathbf{r}_{C_j,3}^i}{\mathbf{r}_{C_j,1}^i} \end{bmatrix}. \tag{3.14}$$

The Kalman filtering approach assumes that measurement noise, $\mathbf{v}_j(k)$, in (3.12) is a discrete-time, zero-mean, white Gaussian process with known covariance, $\mathbf{R}_j(k)$. The proportionality factor L_j in (3.14) represents the effective focal length of the jth camera.

The overall measurement model is written as

$$\mathbf{z}(k) = \mathbf{h}[\mathbf{x}(k)] + \mathbf{v}(k), \tag{3.15}$$

where

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \vdots \\ \mathbf{z}_{N_C} \end{bmatrix}, \ \mathbf{h} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_{N_C} \end{bmatrix}, \text{ and } \mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_{N_C} \end{bmatrix}. \tag{3.16}$$

Implementation of the extended Kalman filter, which is reviewed in Appendix A, and estimation of Cramer Rao bounds, which are reviewed in Appendix B, requires the computation of $\mathbf{r}_{C_j}^i$, $\mathbf{I}_E^O(\zeta)$ and $\mathbf{h}[\mathbf{x}]$ as well as the Jacobian of $\mathbf{h}[\mathbf{x}]$,

$$\mathbf{H}[\mathbf{x}] \triangleq \frac{\partial \mathbf{h}[\mathbf{x}]}{\partial \mathbf{x}}, \tag{3.17}$$

from the most recent state estimate. Computation of $\mathbf{H}[\mathbf{x}]$ is a straight-forward but tedious exercise in differentiation, however, with algebraic manipulation and simplification, $\mathbf{H}$ can be written in simple form which admits computationally efficient implementation. Equations of the measurement model Jacobian for the three parameterizations of rotational motion are given in Appendix D.

3.4 Translational Motion

Translational motion in object tracking problems is commonly modelled with a truncated Taylor series expansion in time. In this case the position, $\mathbf{p}(t) \triangleq [\mathbf{T}_O^E]_E(t)$, of the object-centred frame over any sample period, $\{t |\ t_k \leq t \leq t_{k+1}\}$, is modelled with an $(N_p - 1)$th-order polynomial in time,

$$\mathbf{p}(t) = \sum_{j=0}^{N_p-1} \frac{1}{j!} \frac{\partial^j \mathbf{p}}{\partial t^j}(t_k)[t - t_k]^j. \tag{3.18}$$

A state vector corresponding to translational motion, $\mathbf{x}_t \in \mathbb{R}^{N_t}$, where $N_t = 3N_p$, is defined as

$$\mathbf{x}_t = \left[\mathbf{p}^T, \dot{\mathbf{p}}^T, \ddot{\mathbf{p}}^T, \ldots, \left(\frac{\partial^{N_p-1}\mathbf{p}}{\partial t^{N_p-1}}\right)^T\right]^T, \tag{3.19}$$

which propagates in time according to

$$\dot{\mathbf{x}}_t(t) = \mathbf{A}_t\mathbf{x}_t(t) + \mathbf{G}_t\mathbf{w}_t(t), \tag{3.20}$$

where

$$\mathbf{A}_t = \begin{bmatrix} \mathbf{0}_{3(N_p-1)\times 3} & \mathbf{I}_{3(N_p-1)} \\ \mathbf{0}_{3\times 3(N_p-1)} & \mathbf{0}_{3\times 3} \end{bmatrix}, \text{ and } \mathbf{G}_t = \begin{bmatrix} \mathbf{0}_{3(N_p-1)\times 3} \\ \mathbf{I}_3 \end{bmatrix}, \tag{3.21}$$

and $\mathbf{w}_t \in \mathbb{R}^3$ is a zero mean, temporally white Gaussian process with covariance $\mathbf{Q}_t$.

In this case, the *manoeuvre variable*, $\mathbf{m}(t)$, defined as

$$\mathbf{m}(t) \triangleq \frac{\partial^{N_p}\mathbf{p}}{\partial t^{N_p}}(t) = \mathbf{w}_t(t) \tag{3.22}$$

is modelled as being uncorrelated in time. However, if an object is executing a manoeuvre at time t, it is likely to be executing a similar or perhaps even the same manoeuvre at time $t+\tau$ for sufficiently small τ which implies that $\mathbf{m}(t)$ should be modelled as a time-correlated process. One approach, [27, 71, 75] is to model each component of $\mathbf{m(t)} = [m_1(t), m_2(t), m_3(t)]^T$ as a first-order Gauss-Markov process with autocorrelation function $r_i(\tau)$ given by

$$r_i(\tau) \triangleq \mathcal{E}\{m_i(t)m_i(t+\tau)\} = \sigma_{t_i}^2 \mathrm{e}^{-\alpha_{t_i}|\tau|}, \tag{3.23}$$

where $\sigma_{t_i}^2$ is the variance of the ith component of the manoeuvre variable, and $\alpha_{t_i} \geq 0$ is the reciprocal of the manoeuvre correlation time constant.

The subscripts i on α_{t_i} and σ_{t_i} will be omitted since, in the absence of alternative information, each component of $\mathbf{m}(t)$ can be treated identically. The Wiener-Kalmogorov whitening procedure can then be used to express the manoeuvre variables $m_i(t)$ in terms of white noise by taking the Laplace transform of $r_i(\tau)$ and writing

$$\begin{aligned} R_i(s) &\triangleq \mathcal{L}\{r_i(\tau)\} \\ &= \frac{-2\alpha_t\sigma_t^2}{(s-\alpha_t)(s+\alpha_t)} \\ &\triangleq H(s)H(-s)W_i(s). \end{aligned} \tag{3.24}$$

The function

$$H(s) \triangleq \frac{1}{s+\alpha_t} \tag{3.25}$$

is identified as the Laplace transform of whitening filters for each $m_i(t)$, while

$$W_i(s) \triangleq 2\alpha_t\sigma_t^2 \tag{3.26}$$

represents the power spectral density of independent white noise processes $w_i(t)$ that drive identical filters $H(s)$ to produce $m_i(t)$, according to first-order differential equations

$$\dot{m}_i(t) = -\alpha_t m_i(t) + w_i(t). \tag{3.27}$$

With this model for manoeuvring motion, the state vector $\mathbf{x}_t$ in (3.19) is augmented to include $\mathbf{m}(t)$ of (3.22),

$$\mathbf{x}_t = \left[\mathbf{p}^T, \dot{\mathbf{p}}^T, \ddot{\mathbf{p}}^T, \ldots, \left(\frac{\partial^{N_p}\mathbf{p}}{\partial t^{N_p}}\right)^T\right]^T, \tag{3.28}$$

and propagates in time according to (3.20), but now

$$\mathbf{A}_t = \begin{bmatrix} \mathbf{0}_{3N_p\times 3} & \mathbf{I}_{3N_p} \\ \mathbf{0}_{3\times 3N_p} & -\alpha_t\mathbf{I}_3 \end{bmatrix}, \text{ and } \mathbf{G}_t = \begin{bmatrix} \mathbf{0}_{3N_p\times 3} \\ \mathbf{I}_3 \end{bmatrix}, \tag{3.29}$$

and $\mathbf{w}_t(t)$ is zero-mean, temporally white Gaussian noise with covariance $\mathbf{Q}_t = 2\alpha_t\sigma_t^2\mathbf{I}_3$.

The linear differential equation in (3.20), with the definitions in (3.29), can be written as a discrete-time difference equation using standard techniques to give

$$\mathbf{x}_t(k+1) = \mathbf{\Phi}_t(k)\mathbf{x}_t(k) + \mathbf{w}_t^d(k), \tag{3.30}$$

where

$$\mathbf{\Phi}_t(k) = e^{\mathbf{A}_t(t_{k+1}-t_k)}. \tag{3.31}$$

The state transition matrix, $\mathbf{\Phi}_t$, can be written in closed form due to the simple form of $\mathbf{A}_t$, [71]. The system is driven by a zero-mean, discrete-time, white noise sequence, $\mathbf{w}_t^d(k)$, with covariance $\mathbf{Q}_t^d$ which may be computed (see [71]) directly from Equation (B.35) of Appendix B, or approximated, assuming the sample period is much smaller than system time constants ($2\alpha T \ll 1$, [71]) as

$$\mathbf{Q}_t^d \approx 2\alpha_t\sigma_t^2 T\mathbf{G}_t\mathbf{G}_t^T, \tag{3.32}$$

where T is the sample period (superscript 'T' still denotes transposition).

Three parameters, N_p, α_t, and σ_t, must be specified to complete the translational motion model. Extensive simulations have demonstrated that $N_p = 2$ provides a reasonable balance between accuracy and computational requirements. In fact, because measurements are related directly only to object position, estimation of higher time derivatives ($N_p > 2$) often demonstrated serious lag and overshoot which can lead to filter instability. The parameter α_t is the reciprocal of the manoeuvre correlation time constant. If manoeuvres are expected to occur only over short time periods, α_t should be chosen fairly large (≥ 1), whereas slower expected manoeuvres are modelled with a much smaller α_t ($\approx 1/60$). In the selection of σ_t^2, Singer [71] proposed a model

for the probability distribution of each $m_i(t)$: the maximum and minimum values, $M_{\max}$ and $-M_{\max}$, respectively, each occur with probability $P_{\max}$; $m_i = 0$ occurs with probability P_0; and m_i takes on values in the range $(M_{\max}, -M_{\max})$ according to the appropriate uniform distribution. In this case,

$$\sigma_t^2 = \frac{M_{\max}^2}{3}[1 + 4P_{\max} - P_0]. \tag{3.33}$$

Although this discussion may provide some guidance in the selection of parameters, α_t and σ_t^2 in particular can be treated as tuning parameters in the filtering equations. The first-order Gauss Markov model is also used in the parameterizations of rotational motion developed in the following sections.

3.5 Rotational Motion

The object's instantaneous orientation with respect to the earth-fixed frame is given by the basis transformation $\mathbf{I}_E^O(\zeta)$, where elements of the vector ζ depend on the parameterization selected for rotational motion. The orthogonal transformation $\mathbf{I}_E^O$ belongs to the three-dimensional rotation group for which there are about eight commonly used parameterizations [67]. Appendix A provides a review of the Euler angle-axis, quaternion, and roll-pitch-yaw parameterizations and equations which govern their temporal behavior.

A state vector, $\mathbf{x}_r \in \mathbb{R}^{N_r}$, corresponding to rotational motion is defined which contains elements of ζ and additional states which describe the temporal behavior of ζ. The dynamical model for rotational motion is written as

$$\dot{\mathbf{x}}_r(t) = \mathbf{f}_r[\mathbf{x}_r(t)] + \mathbf{G}_r\mathbf{w}_r(t), \tag{3.34}$$

where $\mathbf{f}_r[\cdot]$ is a vector-valued function, and the zero-mean Gaussian white noise process $\mathbf{w}_r(t)$ is mapped to the rotational motion state space through the matrix $\mathbf{G}_r$. In the angle-axis and roll-pitch-yaw parameterizations, approximate temporal behavior results in linear dynamic models, however, the quaternion formulation results in a nonlinear model.

In navigation systems, measurements of angular velocity are often provided by rate gyroscopes. Applications in navigation, therefore, commonly parameterize rotational motion directly in terms of angular velocity. The angular velocity of F_O with respect to F_E expressed in F_E, denoted $[\mathbf{w}_E^O]_E$, is related to $\mathbf{I}_E^O$ and its time derivative according to

$$[\mathbf{w}_E^O]_E^{\times} = \dot{\mathbf{I}}_E^O \left(\mathbf{I}_E^O\right)^T = \dot{\mathbf{I}}_E^O \mathbf{I}_O^E, \tag{3.35}$$

where $[\mathbf{w}]^{\times}$ denotes the matrix cross-product operator corresponding to the vector

$\mathbf{w} = [w_1, w_2, w_3]^T$ and is defined as

$$\mathbf{w}^{\natural} \triangleq \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix}. \tag{3.36}$$

When no confusion can result, the angular velocity will be simply written as $\mathbf{w} \equiv [\mathbf{w}_E^O]_E$.

Target tracking problems, however, are rarely provided with estimates of angular velocity, but instead must estimate $\mathbf{w}(t)$ based on angular position information embedded in the available measurements. Only the quaternion parameterization is formulated directly in terms of the angular velocity; dynamical models for angle-axis and roll-pitch-yaw parameterizations employ time derivatives of rotational parameters which, in turn, can be used to compute a suboptimal estimate of angular velocity. Even in the quaternion formulation, an extended Kalman filter must be employed and hence suboptimal performance can be expected. Sections 3.5.1, 3.5.2, and 3.5.3 present and discuss dynamical models for the angle-axis, quaternion, and roll-pitch-yaw filters, respectively.

3.5.1 Angle-Axis Parameterization

An orthogonal basis transformation in $\mathbb{R}^3$ can be defined by an *orientation vector* $\xi_E = [\xi_1, \xi_2, \xi_3]_E^T$ whose direction,

$$\bar{\xi}_E \triangleq \frac{\xi_E}{\|\xi_E\|}, \tag{3.37}$$

specifies the axis of the rotation operation relating the two reference frames, and magnitude,

$$\gamma \triangleq \|\xi_E\|, \tag{3.38}$$

specifies the angle through which the coordinate system is rotated about ξ_E, with the direction of rotation taken by the right-hand rule. Since elements of ξ_E are invariant under this rotation operation, the subscript 'E' will be omitted. The basis transformation in this case is computed as

$$\begin{aligned} \mathbf{I}_E^O &= \exp\left(\xi^{\natural}\right) \\ &= \mathbf{I}_3 + \sin(\gamma)\bar{\xi}^{\natural} + [1 - \cos(\gamma)]\left(\bar{\xi}^{\natural}\right)^2, \end{aligned} \tag{3.39}$$

where $\xi^{\natural}$ is the matrix cross product operator, defined in (3.36), associated with ξ.

The axis, $\bar{\xi}$, of $\mathbf{I}_E^O$ is generally distinct from what is commonly called the "axis of rotation" or instantaneous angular velocity[1],

$$\mathbf{w} = \dot{\gamma}\bar{\xi} + \sin\gamma\dot{\bar{\xi}} + (1 - \cos\gamma)\bar{\xi}^{\natural}\dot{\bar{\xi}}, \tag{3.40}$$

[1]All time derivatives are relative to an observer in the earth-fixed frame, and derivative vectors are expressed in the earth-fixed frame.

which has components on three orthogonal axis $\bar{\xi}$, $\dot{\bar{\xi}}$, and $\bar{\xi}^{\parallel}\dot{\bar{\xi}}$ If the axis $\bar{\xi}$ is fixed, (3.40) reduces to the simple expression $\mathbf{w} = \dot{\gamma}\bar{\xi}$. However, even if the object has constant angular velocity, the axis of $\mathbf{I}_E^O$ need not remain fixed.

Under an assumption of "smooth" object motion, approximate temporal behavior of the orientation vector ξ over any sample period, $\{t|t_k \leq t \leq t_{k+1}\}$, is expressed in terms of a first-order Gauss-Markov process in a manner equivalent to that employed for translational motion in Section 3.4. In this case the rotational motion state vector, $\mathbf{x}_r \in \mathbb{R}^{N_r}$, where $N_r = 3(N_\xi + 1)$, is defined as

$$\mathbf{x}_r \triangleq \left[\xi^T, \dot{\xi}^T, \ldots, \left(\frac{\partial^{N_\xi}\xi}{\partial t^{N_\xi}}\right)^T\right]^T, \tag{3.41}$$

and propagates in time according to

$$\dot{\mathbf{x}}_r(t) = \mathbf{A}_r\mathbf{x}_r(t) + \mathbf{G}_r\mathbf{w}_r(t), \tag{3.42}$$

where

$$\mathbf{A}_r = \begin{bmatrix} \mathbf{0}_{3N_\xi \times 3} & \mathbf{I}_{3N_\xi} \\ \mathbf{0}_{3\times 3N_\xi} & -\alpha_r\mathbf{I}_3 \end{bmatrix}, \text{ and } \mathbf{G}_r = \begin{bmatrix} \mathbf{0}_{3N_\xi \times 3} \\ \mathbf{I}_3 \end{bmatrix}, \tag{3.43}$$

and $\mathbf{w}_r(t)$ is zero-mean, temporally white Gaussian noise with covariance $\mathbf{Q}_r = 2\alpha_r\sigma_r^2\mathbf{I}_3$, and the roles of the reciprocal manoeuvre correlation time constant, α_r, and the variance, σ_r^2, of the manoeuvre variable are equivalent to that described in the development of the translational motion model.

A primary difficulty with this approach is that the orientation vector ξ must be reset, first to maintain a finite magnitude (rotation angle) when the object continually rotates in one direction, and second to yield decreasing Cramer Rao bounds. A trajectory $\xi(t)$ in the region $\pi < \|\xi\| < 2\pi$ is equivalent, in the sense that it produces the same rotation matrix, to a trajectory $\zeta(t)$ in the region $\|\zeta\| < \pi$ if

$$\zeta(t) = (\|\xi(t)\| - 2\pi)\bar{\xi}(t) = \xi(t) - 2\pi\bar{\xi}(t). \tag{3.44}$$

This transformation represents a reflection through the origin of $\mathbb{R}^3$, followed by a reflection in the plane tangent to the sphere of radius π at the point $-\pi\bar{\xi}$. One difficulty is that if $\bar{\xi}$ is non-constant, then $\zeta(t)$ may have an infinite number of non-zero time derivatives. Let the orientation vector immediately before the reset be denoted by ξ, and the corresponding orientation vector immediately following the reset is denoted by ζ. An approximate impulsive reset which limits γ to the range

$[0, \pi]$ can be applied according to the algorithm (given only to the first 2 derivatives)

$$
\begin{aligned}
&\text{If } (\|\xi\| > \pi) \text{ then} \\
&\quad \gamma = \|\xi\| \ \{\text{Preliminary Calculations}\} \\
&\quad \bar{\xi} = \xi/\gamma \\
&\quad \dot{\gamma} = \bar{\xi}^T\dot{\xi} \\
&\quad \dot{\bar{\xi}} = (\mathbf{I}_3 - \bar{\xi}\bar{\xi}^T)\dot{\xi}/\gamma \\
&\quad \ddot{\gamma} = \bar{\xi}^T\ddot{\xi} + \dot{\xi}^T\dot{\xi}/\gamma - (\dot{\gamma}/\gamma)^2 \\
&\quad \ddot{\bar{\xi}} = (\mathbf{I}_3 - \bar{\xi}\bar{\xi}^T)(\ddot{\xi}/\gamma - \dot{\xi}\dot{\gamma}/\gamma^2) - (\dot{\bar{\xi}}\bar{\xi}^T + \bar{\xi}\dot{\bar{\xi}}^T)\dot{\xi}/\gamma \\
&\quad \vdots \\
&\quad \zeta = \xi - 2\pi\bar{\xi} \ \{\text{Reset Procedure}\} \\
&\quad \dot{\zeta} = \dot{\xi} - 2\pi\dot{\bar{\xi}} \\
&\quad \ddot{\zeta} = \ddot{\xi} - 2\pi\ddot{\bar{\xi}} \\
&\quad \vdots
\end{aligned}
\tag{3.45}
$$

where, as indicated, the approximate reset can be extended to further time derivatives by differentiating

$$\zeta = (\gamma - 2\pi)\bar{\xi} = \xi - 2\pi\bar{\xi}. \tag{3.46}$$

This model is completely specified by three parameters, N_ξ, α_r, and σ_r^2, which can be selected based on the discussion of translational motion parameters in Section 3.4. Further discussions of the reset and expected performance of the angle-axis filter are deferred to Chapter 4 which compares Cramer Rao bounds and presents simulation results.

3.5.2 Quaternion Parameterization

A quaternion [29]-[37], [65]-[70], [83], [84], is a four dimensional hypercomplex number[2]

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = q_1 i + q_2 j + q_3 k + q_4, \tag{3.47}$$

which is expressed in terms of basis elements consisting of the real number $+1$, and three imaginary units i, j, k, which satisfy

$$
\begin{aligned}
&i^2 = j^2 = k^2 = -1, \\
&ij = -ji = k, \\
&jk = -kj = i, \quad \text{and} \\
&ki = -ik = j.
\end{aligned}
\tag{3.48}
$$

[2]Some authors place the scalar quaternion element, q_4, in the first position while others place the scalar element in the fourth position, as is the case here, when representing the quaternion in vector form.

If $\mathbf{I}_E^O(\mathbf{q})$ rotates $\mathbb{R}^3$ through an angle γ about the axis $\bar{\xi} = [\bar{\xi}_1, \bar{\xi}_2, \bar{\xi}_3]^T$ to bring F_E into alignment with F_O, the unit quaternion defined by

$$\begin{aligned} \bar{\mathbf{q}} &\triangleq \frac{\mathbf{q}}{\|\mathbf{q}\|} \\ &\triangleq \left[-\bar{\xi}_1 \sin(\frac{\gamma}{2}), -\bar{\xi}_2 \sin(\frac{\gamma}{2}), -\bar{\xi}_3 \sin(\frac{\gamma}{2}), \cos(\frac{\gamma}{2})\right]^T \end{aligned} \tag{3.49}$$

gives $\mathbf{I}_E^O(\mathbf{q})$ as

$$\mathbf{I}_E^O(\mathbf{q}) = \frac{1}{\|\mathbf{q}\|^2}\mathbf{\Gamma}(\mathbf{q}) = \mathbf{\Gamma}(\bar{\mathbf{q}}), \tag{3.50}$$

where

$$\mathbf{\Gamma}(\mathbf{q}) \triangleq \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_1q_3 - q_2q_4) \\ 2(q_1q_2 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 + q_1q_4) \\ 2(q_1q_3 + q_2q_4) & 2(q_2q_3 - q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix}. \tag{3.51}$$

The matrix $\mathbf{\Gamma}(\mathbf{q})$ represents combined operations of an orthogonal transformation ($\mathbf{I}_E^O$) and uniform scaling by $\|\mathbf{q}\|^2$. The quaternion parameterization is four-dimensional and is often fixed with a quadratic constraint of the form $\|\mathbf{q}\| = 1$ since $\mathbf{I}_E^O(\mathbf{q})$ depends only on the normalized quaternion $\bar{\mathbf{q}}$. Such constraints are not easily incorporated into the Kalman filtering equations due to the linear structure of the filter.

A common approach, [29]-[37], [53], is to implement the filter based on unit quaternions and employ an impulsive normalization of the estimated quaternion, using standard Kalman filtering notation from Appendix B,

$$\hat{\mathbf{q}}(k|k) \leftarrow \frac{\hat{\mathbf{q}}(k|k)}{\|\hat{\mathbf{q}}(k|k)\|}, \tag{3.52}$$

immediately following each observation event. In the present work, this approach resulted in strict and total divergence of angular velocity estimates in a significant portion (sometimes more than 80%) of simulation runs. Tahk and Spayer [66] propose to incorporate soft or stochastic constraints directly into the measurement model and apply the extended Kalman filter to the augmented system. This approach, although it resulted in an improvement over impulsive normalization, also led to poor and often divergent behavior in angular velocity estimation. Bar-Itzack and Oshman [65] treat the problem of estimating, in the current notation, a unit quaternion $\bar{\mathbf{q}}$, and hence $\mathbf{I}_E^O(\bar{\mathbf{q}})$, from sequences of observations of corresponding vectors $\{(\mathbf{r}_E^i, \mathbf{r}_C^i); i = 1, 2, \ldots, N\}$, and angular velocity $\mathbf{w}$ over time with an extended Kalman filter. With a first-order approximation, their approach is to propagate the estimated quaternion in the filter without impulsive normalization, which means that the estimated quaternion need not have unit norm. For output purposes only, the estimated quaternion is normalized through division by its Euclidean norm [64]. Their problem is slightly different from the one presently under consideration, but the main ideas of their approach have been employed in the quaternion filter with slight modifications.

The dynamical model is based on the assumption that the quaternion has constant norm (not necessarily unity) so that

$$\dot{\mathbf{q}}(t) = \|\mathbf{q}(t)\|\dot{\bar{\mathbf{q}}}(t). \tag{3.53}$$

The unit quaternion, $\bar{\mathbf{q}}(t)$ defined in (3.49), propagates in time (see Appendix A) according to the differential equation

$$\dot{\bar{\mathbf{q}}}(t) = \boldsymbol{\Omega}[\mathbf{w}(t)]\bar{\mathbf{q}}(t), \tag{3.54}$$

where $\mathbf{w} \equiv [\mathbf{w}_E^O]_E = [w_1, w_2, w_3]^T$ and

$$\boldsymbol{\Omega}[\mathbf{w}] \triangleq \frac{1}{2}\begin{bmatrix} 0 & -w_3 & w_2 & -w_1 \\ w_3 & 0 & -w_1 & -w_2 \\ -w_2 & w_1 & 0 & -w_3 \\ w_1 & w_2 & w_3 & 0 \end{bmatrix}. \tag{3.55}$$

Multiplying both sides of (3.54) by $\|\mathbf{q}\|$ gives, with (3.53),

$$\dot{\mathbf{q}}(t) = \boldsymbol{\Omega}[\mathbf{w}(t)]\mathbf{q}(t). \tag{3.56}$$

Once again employing the assumption that rotational motion is "smooth", temporal behavior of $\mathbf{w}$ over any sample period is approximated by a first-order Gauss-Markov process in a manner equivalent to that employed for translational motion in Section 3.4. A state vector, $\mathbf{x}_w$, for the angular velocity and its time derivatives is defined as

$$\mathbf{x}_w \triangleq \left[\mathbf{w}^T, \dot{\mathbf{w}}^T, \ldots, \left(\frac{\partial^{N_w}\mathbf{w}}{\partial t^{N_w}}\right)^T\right]^T, \tag{3.57}$$

which propagates in time according to

$$\dot{\mathbf{x}}_w(t) = \mathbf{A}_w\mathbf{x}_w(t) + \mathbf{G}_w\mathbf{w}_r(t), \tag{3.58}$$

where

$$\mathbf{A}_w = \begin{bmatrix} \mathbf{0}_{3N_w\times 3} & \mathbf{I}_{3N_w} \\ \mathbf{0}_{3\times 3N_w} & -\alpha_r\mathbf{I}_3 \end{bmatrix}, \text{ and } \mathbf{G}_w = \begin{bmatrix} \mathbf{0}_{3N_w\times 3} \\ \mathbf{I}_3 \end{bmatrix}, \tag{3.59}$$

and $\mathbf{w}_r(t)$ is zero-mean, temporally white Gaussian noise with covariance $\mathbf{Q}_r = 2\alpha_r\sigma_r^2\mathbf{I}_3$, and the roles of the reciprocal manoeuvre correlation time constant, α_r, and the variance, σ_r^2, of the manoeuvre variable are equivalent to that described in the development of the translational motion model.

An overall state vector, $\mathbf{x}_r$, corresponding to rotational motion is defined as

$$\mathbf{x}_r \triangleq \left[\mathbf{q}^T, \mathbf{x}_w^T\right]^T, \tag{3.60}$$

which propagates in time according to the nonlinear differential equation

$$\begin{aligned}\dot{\mathbf{x}}_r(t) &= \mathbf{f}_r[\mathbf{x}_r(t)] + \mathbf{G}_r\mathbf{w}_r(t) \\ &= \begin{bmatrix} \boldsymbol{\Omega}[\mathbf{w}(t)] & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_w \end{bmatrix} \mathbf{x}_r(t) + \begin{bmatrix} \mathbf{0}_{4\times 3} \\ \mathbf{G}_w \end{bmatrix} \mathbf{w}_r(t). \end{aligned} \tag{3.61}$$

Time propagation in the extended Kalman filter requires numerical integration of (3.61) with $\mathbf{G}_r\mathbf{w}_r(t) = \mathbf{0}$.

If $\mathbf{w}$ is approximately constant over any sample period, an approximate discrete-time difference equation can be obtained for $\mathbf{x}_r$ of the form

$$\mathbf{x}_r(k+1) = \boldsymbol{\Phi}_r(k)\mathbf{x}_r(k), \tag{3.62}$$

where

$$\boldsymbol{\Phi}_r(k) \triangleq \text{diag}\left[\boldsymbol{\Phi}_q(k), \boldsymbol{\Phi}_w(k)\right], \tag{3.63}$$

and

$$\boldsymbol{\Phi}_q(k) \triangleq \exp(\boldsymbol{\Omega}[\mathbf{w}(k)]\{t_{k+1} - t_k\}) \tag{3.64}$$

$$\boldsymbol{\Phi}_w(k) \triangleq \exp(\mathbf{A}_w\{t_{k+1} - t_k\}). \tag{3.65}$$

A closed-form expression for $\boldsymbol{\Phi}_q$ (see Appendix A) is

$$\boldsymbol{\Phi}_q(k) = \cos\left(\|\mathbf{w}\|\frac{(t_{k+1}-t_k)}{2}\right)\mathbf{I}_4 + \sin\left(\|\mathbf{w}\|\frac{(t_{k+1}-t_k)}{2}\right)\frac{2}{\|\mathbf{w}\|}\boldsymbol{\Omega}[\mathbf{w}]. \tag{3.66}$$

Because $\boldsymbol{\Omega}[\mathbf{w}]$ is a skew symmetric matrix, $\boldsymbol{\Phi}_q(k)$ is orthogonal[3] and hence $\|\mathbf{q}(k+1)\| = \|\mathbf{q}(k)\|$ as required.

Numerical integration in (3.62) is based on the assumption that $\mathbf{w}$ is approximately constant over the sample period. This will be essentially true provided either $\mathbf{w}$ is constant, or the sample period is very small. An assumption of small sample periods is often employed in motion analysis, but is very restrictive. Moreover, this approximation does not exploit estimates of time derivatives of $\mathbf{w}$. Many standard methods for numerical integration of the quaternion differential equation are designed for navigation systems in which measurements of angular velocity are available directly [84]. The following algorithm is similar in form to one analyzed by Branets and Shmyglevski [83]. This approach employs the modelled temporal behavior of $\mathbf{w}$ and the most recent estimates of $\mathbf{w}$ and its time derivatives.

Define a partition

$$\Pi(k) \triangleq \{t_k = \tau_0 \leq \tau_1 \leq \cdots \tau_{N_\pi} = t_{k+1}\} \tag{3.67}$$

[3]The matrix exponential of a skew symmetric matrix is always orthogonal.

of the kth sample period, and assume for convenience that $\tau_{\ell+1} - \tau_\ell \triangleq \tau$ is constant. With the linear model for $\mathbf{x}_w$,

$$\begin{aligned} \mathbf{x}_w(\tau_\ell) &= [\Phi_w(\tau)]^\ell \mathbf{x}_w(t_k) \\ \mathbf{w}(\tau_\ell) &= \mathbf{C}_w \mathbf{x}_w(\tau_\ell) \end{aligned} \tag{3.68}$$

where

$$\begin{aligned} \Phi_w(\tau) &\triangleq e^{\mathbf{A}_w \tau}, \text{ and} \\ \mathbf{C}_w &\triangleq \begin{bmatrix} 1,0,0,0,\ldots,0 \\ 0,1,0,0,\ldots,0 \\ 0,0,1,0,\ldots,0 \end{bmatrix}. \end{aligned} \tag{3.69}$$

Provided τ is chosen small enough, or that N_π is chosen large enough, so that $\mathbf{w}$ is approximately constant over any interval of the partition $\Pi(k)$, time propagation of $\mathbf{q}$ over the complete sample period may be approximated by

$$\mathbf{q}(k+1) = \Phi_q(k)\mathbf{q}(k), \tag{3.70}$$

where

$$\Phi_q(k) \triangleq \downarrow \prod_{\ell=0}^{N_\pi - 1} \exp\left\{\Omega\left[\mathbf{C}_w \left(\Phi_w(\tau)\right)^\ell \mathbf{x}_w(t_k)\right]\tau\right\}, \tag{3.71}$$

in which the notation $\downarrow \prod$ indicates that the index ℓ decreases from left to right in the matrix product, $\Omega[\cdot] : \mathbb{R}^3 \to \mathbb{R}^{4\times 4}$ according to (3.55), and (3.66) can be used to evaluate the matrix exponentials. Note that Φ_q in (3.71) is a product of orthogonal matrices and hence is orthogonal, which means that $\|\mathbf{q}(k+1)\| = \|\mathbf{q}(k)\|$.

It should be noted that the continuous-time form, $\Phi_r(t,\lambda)$, of $\Phi_r(k)$ obtained by replacing t_{k+1} with t and t_k by λ in (3.64) or (3.71) and (3.65), is not a "state transition matrix" in the usual sense because it does **not** satisfy the fundamental property

$$\frac{d\Phi_r(t,\lambda)}{dt} = \frac{\partial \mathbf{f}_r[\mathbf{x}_r]}{\partial \mathbf{x}_r}\Phi_r(t,\lambda), \quad \text{(false)}. \tag{3.72}$$

As a result, even though Φ_r may be used for approximate time propagation of the state estimate, this matrix cannot be used to propagate the Kalman covariance matrix over time. At any measurement event, the measurements are directly related to the quaternions and not to the angular velocity. The Kalman filter must use the dynamical model to extract information about angular velocity from prediction errors which, in turn, depend on the quaternion estimates. Time propagation of the Kalman covariance, therefore, must maintain the interaction between quaternions and the angular velocity specified in (3.56). The block diagonal matrix Φ_r in (3.63) does not explicitly model the interdependence of $\mathbf{q}$ and $\mathbf{w}$. Approximate time propagation, [29]-[37], [53], of the covariance matrix is performed with an approximate state transition matrix which employs the rotational motion component as

$$\hat{\Phi}_r(k) = \exp\{\mathbf{F}_r(k)[t_{k+1} - t_k]\}, \tag{3.73}$$

where $\mathbf{F}_r(k)$ is the Jacobian of $\mathbf{f}_r[\mathbf{x}_r(k)]$,

$$\begin{aligned} \mathbf{F}_r(k) &\triangleq \frac{\partial \mathbf{f}_r[\mathbf{x}_r]}{\partial \mathbf{x}_r} \\ &\triangleq \begin{bmatrix} \boldsymbol{\Omega}[\mathbf{w}(k)] & [\dot{\mathbf{q}}_w(k) \ \mathbf{0}] \\ \mathbf{0} & \mathbf{A}_w \end{bmatrix}, \end{aligned} \tag{3.74}$$

and

$$\dot{\mathbf{q}}_w(k) \triangleq \frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{w}}(k) = \frac{1}{2}\begin{bmatrix} -q_4 & q_3 & -q_2 \\ -q_3 & -q_4 & q_1 \\ q_2 & -q_1 & -q_4 \\ q_1 & q_2 & q_3 \end{bmatrix}. \tag{3.75}$$

Note that $\hat{\boldsymbol{\Phi}}_r(k)$ in (3.73) need not in general maintain constant quaternion norm and hence should not be used in time propagation of the state estimate.

The quaternion filter has been implemented, using standard Kalman filtering notation introduced in Appendix B, with the following approach:

1. The filter is initialized with a unit quaternion, $\|\hat{\mathbf{q}}(0|-1)\| = 1$;
2. The norm of the estimated quaternion is invariant under time propagation, i.e., $\|\hat{\mathbf{q}}(k+1|k)\| = \|\hat{\mathbf{q}}(k|k)\|$, $k = 1, 2, 3, \ldots$;
3. The measurement model employs $\boldsymbol{\Gamma}(\hat{\mathbf{q}})$ of (3.51) as the transformation from F_O to F_E instead of the strictly orthogonal matrix $\mathbf{I}_E^O$;
4. For output purposes only, the estimated unit quaternion is taken as

$$\bar{\hat{\mathbf{q}}}(k|k) \triangleq \frac{\hat{\mathbf{q}}(k|k)}{\|\hat{\mathbf{q}}(k|k)\|}, \tag{3.76}$$

and, again for output purposes only, structure estimates denoted ${}^*\hat{\mathbf{r}}_O^i(k|k)$ are taken as

$${}^*\hat{\mathbf{r}}_O^i(k|k) \triangleq \|\hat{\mathbf{q}}(k|k)\|^2 \hat{\mathbf{r}}_O^i(k|k), \tag{3.77}$$

for $i = 1, 2, \ldots, N_f$, where the vectors $\hat{\mathbf{r}}_O^i(k|k)$ are structure estimates generated within the filter;

5. The estimation error covariance matrix, $\hat{\mathbf{P}}^*(k|k)$, for the filter output is taken as

$$\hat{\mathbf{P}}^*(k|k) \triangleq \boldsymbol{\beta}\hat{\mathbf{P}}(k|k)\boldsymbol{\beta}^T, \tag{3.78}$$

where $\hat{\mathbf{P}}(k|k)$ is the estimated Kalman covariance, and

$$\boldsymbol{\beta} \triangleq \operatorname{diag}\left\{\mathbf{I}_{N_t}, \frac{1}{\|\hat{\mathbf{q}}(k|k)\|}\mathbf{L}_4, \mathbf{I}_{N_w}, \|\hat{\mathbf{q}}(k|k)\|^2 \mathbf{I}_{3N_f}\right\}. \tag{3.79}$$

This approach parallels that proposed by Bar-Itzhack and Oshman [65]. Due to Item 3 above, the measurement model depends on the product of $\Gamma(q)$ and structure vectors $\mathbf{r}_O^i$ internal to the filter which appears in the expression,

$$\mathbf{r}_E^i = \Gamma(q)\mathbf{r}_O^i + [\mathbf{T}_E^O]_E. \tag{3.80}$$

The innovation sequence in the quaternion filter, therefore, depends on the product

$$\widehat{[\mathbf{r}_O^i]_E} \triangleq \Gamma(\hat{\mathbf{q}})\hat{\mathbf{r}}_O^i. \tag{3.81}$$

If the quaternion filter works well, $\widehat{[\mathbf{r}_O^i]_E} \rightarrow [\mathbf{r}_O^i]_E$ for $i = 1, 2, \ldots, N_f$, even though $\hat{\mathbf{q}}$ may not have unit norm, and $\hat{\mathbf{r}}_O^i$ may not converge to a true structure vector (due to the scaling effect of $\Gamma(\mathbf{q})$). Noting from (3.50) and (3.51) that

$$\begin{aligned} \widehat{[\mathbf{r}_O^i]_E} &\triangleq \Gamma(\hat{\mathbf{q}})\hat{\mathbf{r}}_O^i \\ &= \left\{\frac{1}{\|\hat{\mathbf{q}}\|^2}\Gamma(\hat{\mathbf{q}})\right\}\left\{\|\hat{\mathbf{q}}\|^2\hat{\mathbf{r}}_O^i\right\} \\ &= \{\mathbf{I}_E^O(\bar{\hat{\mathbf{q}}})\}\{{}^*\hat{\mathbf{r}}_O^i(k|k)\} \end{aligned} \tag{3.82}$$

leads naturally to the extraction, in item 4 above, of a unit quaternion and appropriate structure estimates.

This quaternion model is completely specified by four parameters, N_w, α_r, σ_r^2, and N_{pi}. The first three of these can be selected based on the discussion of translational motion parameters in Section 3.4. Selection of number of intervals, N_π, in the partition of each sample period should also be based in part on the expected complexity of rotational motion. Investigations leading to the research reported herein indicated that $N_\pi = 5$ with a sample period of $T = 0.1$ seconds gives reasonable accuracy in ideal time propagation of the state at the expense of increased computational complexity.

3.5.3 Roll Pitch Yaw Parameterization

The roll-pitch-yaw (RPY) representation [38, 63, 82] is a particular case of an Euler angle parameterization of the three-dimensional rotation group. The roll, pitch, and yaw angles, denoted ϕ, θ, and ψ, respectively, define an ordered sequence of three plane rotations which can be used to express $\mathbf{I}_E^O$ as

$$\mathbf{I}_E^O = \exp(\psi\mathbf{e}_3')\exp(\theta\mathbf{e}_2')\exp(\phi\mathbf{e}_1'), \tag{3.83}$$

where $\mathbf{e}_i$, $i = 1, 2, 3$, are the standard basis vectors of $\mathbb{R}^3$. Expanding the exponentials with (A.13) gives three plane rotations,

$$\exp(\phi\mathbf{e}_1') = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi) & \cos(\phi) \end{bmatrix},$$

$$\exp(\theta \mathbf{e}_2^{\natural}) = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix}, \text{ and} \tag{3.84}$$

$$\exp(\psi \mathbf{e}_3^{\natural}) = \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The angles ϕ, θ, and ψ have positive sense defined by the right hand rule about their respective rotation axis (see Appendix A).

It has been reported, [28, 32, 37], that major difficulties arise with this approach because the sequence of plane rotations are not defined about orthogonal axis. This property results in highly nonlinear dynamical models if rotational motion is formulated directly in terms of angular velocity. The angular velocity $\mathbf{w} \equiv [\mathbf{w}_E^O]_E$, is given in terms of time derivatives of roll, pitch, and yaw angles as

$$\mathbf{w} = \mathbf{J}_{RPY} \begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix}, \tag{3.85}$$

where

$$\mathbf{J}_{RPY} = \begin{bmatrix} \cos(\psi)\cos(\theta) & -\sin(\psi) & 0 \\ \sin(\psi)\cos(\theta) & \cos(\psi) & 0 \\ -\sin(\theta) & 0 & 1 \end{bmatrix}. \tag{3.86}$$

The mathematical singularities, $\theta = \pm\pi/2$, characterize orientations where the determinant of the Jacobian matrix $\det[J_{RPY}] = -\cos(\theta)$ vanishes.

The parameter vector for the roll-pitch-yaw filter is defined as

$$\zeta \triangleq [\phi, \theta, \psi]^T. \tag{3.87}$$

Under an assumption of "smooth" object motion, an approximate temporal behavior of ζ over any sample period, $\{t | t_k \leq t \leq t_{k+1}\}$, is expressed in terms of a first-order Gauss-Markov process in a manner equivalent to that employed for translational motion in Section 3.4. In this case the rotational motion state vector, $\mathbf{x}_r \in \mathbb{R}^{N_r}$, where $N_r = 3(N_\zeta + 1)$, is defined as

$$\mathbf{x}_r \triangleq \left[\zeta^T, \dot{\zeta}^T, \ldots, \left(\frac{\partial^{N_\zeta}\zeta}{\partial t^{N_\zeta}}\right)^T\right]^T, \tag{3.88}$$

and propagates in time according to

$$\dot{\mathbf{x}}_r(t) = \mathbf{A}_r \mathbf{x}_r(t) + \mathbf{G}_r \mathbf{w}_r(t), \tag{3.89}$$

where

$$\mathbf{A}_r = \begin{bmatrix} \mathbf{0}_{3N_\zeta \times 3} & \mathbf{I}_{3N_\zeta} \\ \mathbf{0}_{3\times 3N_\zeta} & -\alpha_r \mathbf{I}_3 \end{bmatrix}, \text{ and } \mathbf{G}_r = \begin{bmatrix} \mathbf{0}_{3N_\zeta \times 3} \\ \mathbf{I}_3 \end{bmatrix}, \tag{3.90}$$

and $\mathbf{w}_r(t)$ is zero-mean, temporally white Gaussian noise with covariance $\mathbf{Q}_r = 2\alpha_r\sigma_r^2\mathbf{I}_3$, and the roles of the reciprocal manoeuvre correlation time constant, α_r, and the variance, σ_r^2, of the manoeuvre variable are equivalent to that described in the development of the translational motion model.

The roll-pitch-yaw model is completely specified by three parameters, N_ζ, α_r, and σ_r^2 which, again, can be selected based on the discussion of translational motion parameters in Section 3.4.

3.6 Initial Estimates

The extended Kalman filter requires an initial estimate of the state, $\mathbf{x}(0|-1)$, and characterization of the initial uncertainty, $\mathbf{P}(0|-1)$, in this estimate. Broida and Chellappa, for example, have studied a batch approach [30, 31, 32, 37] which employs the Conjugate Gradient descent technique from the IMSL library. Simulation results shown by Broida *et al.*, [32], employ what are called "crude" initial estimates obtained from the first 10 image frames with 75 iterations of the batch technique.

In future work and eventual applications, it is expected that initialization of object tracking filters would be based on information from lower subsystems in a hierarchical structure such as that described in Section 1.2. For the purpose of this investigation, a simple single-frame initialization provides suitable estimates of position, orientation, and structure, while all higher-order motion parameters (time derivatives) are initialized as zero. The initialization algorithm is described as follows:

Step 1 Obtain an estimate of the 3D positions of three or more feature points with respect to the earth-fixed frame through steriopsis based on noisy measurements in multiple-camera images and assumed correspondence information;

Step 2 Select three noncollinear feature points from all available 3D points;

Step 3 Estimate the initial position of the object, form an orthonormal, right-handed triad, F_O, in F_E from the three special 3D points based on constraints detailed in Section 3.2 and illustrated in Figure 3.2, and fix the origin, O_O, of F_O with respect to the 3D feature points by specifying the vector $\mathbf{d}$;

Step 4 Extract a valid set of orientation parameters from the basis vectors of F_O expressed in F_E;

Step 5 Express positions of all feature points and $\mathbf{d}$ with respect to F_O.

This initialization is concerned only with obtaining an initial estimate of the state; the initial covariance matrix is selected as diagonal assuming significant initialization errors (fairly large diagonal entries).

Step 1: Given $N_f \geq 3$ feature points visible in each of $N_C \geq 2$ cameras with spatial correspondence information, the objective is to estimate $\mathbf{r}^i_E$, $i = 1, 2, \ldots, N_f$. The position of the ith feature point in the image plane of the jth camera is written, from (2.19), as

$$\mathbf{z}^i_j \triangleq \begin{bmatrix} u^i_j \\ v^i_j \end{bmatrix} = L_j \begin{bmatrix} \dfrac{\mathbf{r}^i_{C_j,2}}{\mathbf{r}^i_{C_j,1}} \\ \dfrac{\mathbf{r}^i_{C_j,3}}{\mathbf{r}^i_{C_j,1}} \end{bmatrix}. \tag{3.91}$$

With (3.91) and

$$\mathbf{D}^i_j \triangleq \begin{bmatrix} u^i_j & -L_j & 0 \\ v^i_j & 0 & -L_j \end{bmatrix}, \tag{3.92}$$

so that $\mathbf{D}^i_j \mathbf{r}^i_{C_j} = \mathbf{0}$ and

$$\mathbf{r}^i_{C_j} = \mathbf{I}^E_{C_j} \mathbf{r}^i_E + [\mathbf{T}^{C_j}_E]_{C_j}, \tag{3.93}$$

where $\mathbf{I}^E_{C_j}$ and $[\mathbf{T}^{C_j}_E]_{C_j}$ are known for $j = 1, 2, \ldots, N_C$, a system of linear equations can be written in the form

$$\mathbf{D}^i_j \mathbf{I}^E_{C_j} \mathbf{r}^i_E = -\mathbf{D}^i_j [\mathbf{T}^{C_j}_E]_{C_j}, \quad j = 1, 2, \ldots, N_C. \tag{3.94}$$

Forming two matrices

$$\mathbf{M}^i \triangleq \begin{bmatrix} \mathbf{D}^i_1 \mathbf{I}^E_{C_1} \\ \mathbf{D}^i_2 \mathbf{I}^E_{C_2} \\ \vdots \\ \mathbf{D}^i_{N_C} \mathbf{I}^E_{C_{N_C}} \end{bmatrix}, \text{ and } \mathbf{N}^i \triangleq \begin{bmatrix} \mathbf{D}^i_1 [\mathbf{T}^E_{C_1}]_{C_1} \\ \mathbf{D}^i_2 [\mathbf{T}^E_{C_2}]_{C_2} \\ \vdots \\ \mathbf{D}^i_{N_C} [\mathbf{T}^E_{C_{N_C}}]_{C_{N_C}} \end{bmatrix}, \tag{3.95}$$

immediately gives the least-mean-squared estimate $\hat{\mathbf{r}}^i_E$ of $\mathbf{r}^i_E$ as

$$\hat{\mathbf{r}}^i_E = \left[\left(\mathbf{M}^i\right)^T \mathbf{M}^i \right]^{-1} \left(\mathbf{M}^i\right)^T \mathbf{N}^i \tag{3.96}$$

for the ith feature point. This process is repeated for each of the N_f feature points.

Step 2: Many criteria could be specified for selecting the best set of three special noncollinear feature points from all available $\hat{\mathbf{r}}^i_E$. Moreover, this selection might be aided by decisions from lower-level subsystems of a hierarchical approach such as that discussed in Section 1.2. A simple approach is to select those three points for which the area of the triangle (see Figure 3.2) formed by joining the three 3D points is maximized, thus maximizing a measure of the degree of noncollinearity and distance between points. This can be done by selecting 3 indices, $\{i_1, i_1, i_3\} \subseteq \{1, 2, \ldots, N_f\}$ such that,

$$\| \left[\hat{\mathbf{r}}^{i_2}_E - \hat{\mathbf{r}}^{i_1}_E \right]^{\parallel} \left[\hat{\mathbf{r}}^{i_3}_E - \hat{\mathbf{r}}^{i_1}_E \right] \|^2 \tag{3.97}$$

is maximized. Since the order of selection is unimportant, there are

$$\begin{pmatrix} N_f \\ 3 \end{pmatrix} \triangleq \frac{N_f!}{3!(N_f-3)!} \tag{3.98}$$

combinations for comparison.

Step 3: The estimated initial position, $\hat{\mathbf{p}}(0) \triangleq [\widehat{\mathbf{T}_O^E}]_E(0)$, of the object is taken as the sample mean of all estimated feature point positions:

$$\hat{\mathbf{p}}(0) = \frac{1}{N_f}\sum_{i=1}^{N_f} \hat{\mathbf{r}}_E^i. \tag{3.99}$$

The three feature points selected in Step 2 are ordered, arbitrarily in the absence of prior information, as $\{\hat{\mathbf{r}}_E^{i_0}, \mathbf{r}_E^{i_x}, \hat{\mathbf{r}}_E^{i_{xy}}\}$ according to the constraints detailed in Section 3.2. The object-centred frame is fixed on the object, and the initial estimates of orientation are provided with

$$[\mathbf{d}_O]_E = \hat{\mathbf{r}}_E^{i_0} - \hat{\mathbf{p}}(0), \tag{3.100}$$

$$\widehat{[\mathbf{i}_O]_E} = \frac{\hat{\mathbf{r}}_E^{i_x} - \hat{\mathbf{r}}_E^{i_0}}{\|\hat{\mathbf{r}}_E^{i_x} - \hat{\mathbf{r}}_E^{i_0}\|}, \tag{3.101}$$

$$\widehat{[\mathbf{j}_O]_E} = \frac{[\mathbf{I}_3 - \widehat{[\mathbf{i}_O]_E}(\widehat{[\mathbf{i}_O]_E})^T]\hat{\mathbf{r}}_E^{i_{xy}}}{\|[\mathbf{I}_3 - \widehat{[\mathbf{i}_O]_E}(\widehat{[\mathbf{i}_O]_E})^T]\hat{\mathbf{r}}_E^{i_{xy}}\|}, \text{ and} \tag{3.102}$$

$$\widehat{[\mathbf{k}_O]_E} = \widehat{[\mathbf{i}_O]_E} \times \widehat{[\mathbf{j}_O]_E}. \tag{3.103}$$

Step 4: Initial estimates of orientation parameters (angle-axis, quaternion, or roll-pitch-yaw) are extracted from the initial estimate of the basis transformation

$$\widehat{\mathbf{I}_E^O} = \left[\widehat{[\mathbf{i}_O]_E}, \widehat{[\mathbf{j}_O]_E}, \widehat{[\mathbf{k}_O]_E}\right], \tag{3.104}$$

according to methods outlined in Appendix A.

Step 5: Initial structure estimates with respect to F_O are computed with

$$\mathbf{d}_O = \widehat{\mathbf{I}_E^O}^T [\mathbf{d}_O]_E, \text{ and} \tag{3.105}$$

$$\hat{\mathbf{r}}_O^i = \widehat{\mathbf{I}_E^O}^T \left(\hat{\mathbf{r}}_E^i - \hat{\mathbf{p}}(0)\right), \quad i = 1, 2, \ldots, N_f. \tag{3.106}$$

By construction, $\hat{\mathbf{r}}_O^{i_0} = \mathbf{d}_O$, $\hat{r}_{O,2}^{i_x} = d_{O,2}$, $\hat{r}_{O,3}^{i_x} = d_{O,3}$, and $\hat{r}_{O,3}^{i_{xy}} = d_{O,3}$, where $\mathbf{d}_O$ is now treated as a known constant vector; these six structural parameters are not estimated by the filters.

3.7 Chapter Summary

This chapter has introduced structure, translational motion, rotational motion, and measurement state space models for manoeuvring objects observed with a multiple-camera imaging system. Both translational and rotational motion are described with first-order Gauss-Markov processes which are completely defined by selection of a small number of parameters: the number of time derivatives (N_t and one of N_ξ, N_w, or N_ζ); the reciprocal manoeuvre correlation time constants (α_t and α_r); and the manoeuvre variances (σ_t^2 and σ_r^2). The quaternion filter also requires selection of the number of intervals, N_π, of each sample period for time propagation of the state estimate, and all filters require estimation of the measurement noise covariance, $\mathbf{R}$. A simple single-frame initialization algorithm has been proposed for object position, orientation, and structure. Chapter 4 discusses parameter selection, measures of optimal performance provided by Cramer Rao bounds, and presents simulation results and performance comparisons for filters formed from the three parameterizations of rotational motion.

4. Performance With Simulated Imagery

The previous chapter introduced motion, structure and measurement models based on three parameterizations of rotational motion. In this chapter, relative performance of three iterated extended Kalman filters, each based on a single parameterization, is examined and demonstrated through presentation of a subset of results generated from extensive Monte Carlo simulations. One limitation of analysis based on simulated imagery is that the assumed central projection image formation model is exact; that is, the measurement model is based on the central projection transformation, and simulated imagery is generated precisely from this model. This approach, which is common to most simulation studies in image analysis, is only valid if distortion effects in the imaging system are small in comparison to measurement noise levels. Evaluation of motion and structure estimation schemes on simulated imagery, however, provides an efficient means to verify and compare performance of the three filtering algorithms.

The geometry of Figure 3.1 was used to generate binocular image sequences of a single, rigid, translating and rotating block. Figure 4.1 shows the object used for simulations with the convention for labelling of feature points. Features of interest are four, p^3, p^5, p^6, and p^7, of the eight corners of a rectangular block 20 units wide ($p^6 \rightarrow p^3$) and 10 units in height ($p^6 \rightarrow p^5$) and depth ($p^6 \rightarrow p^7$). Feature detection and both temporal and spatial correspondence information is provided by the imagery generation program, and occlusion is not explicitly considered in this work.

In all simulations of this Chapter, two cameras are employed with unity focal length. The position and orientation of the left camera (C_1) is defined by a translation $[\mathbf{T}^E_{C_1}]_E = [0, -5, -30]^T_E$ units, and transformation $\mathbf{I}^E_{C_1}$ formed from Euler angles $\psi_{C_1} = 10°$ (yaw or pan), $\theta_{C_1} = -45°$ (pitch or tilt), and $\phi_{C_1} = 0$ (roll about optical axis). The position and orientation of the right camera (C_2) is defined, respectively, by $[\mathbf{T}^E_{C_2}]_E = [0, 5, -30]^T_E$ units, and $\mathbf{I}^E_{C_2}$ formed from Euler angles $\psi_{C_2} = -10°$, $\theta_{C_2} = -45°$, and $\phi_{C_2} = 0$. Measurement error in feature-based motion analysis studies is often taken as uniform digitization noise [19, 20, 32, 34]. Simulations of this Chapter assume moderately high levels of measurement noise in the range ± 0.02 units in the image plane, which corresponds to errors on the order of 4% to 8% of the size of the

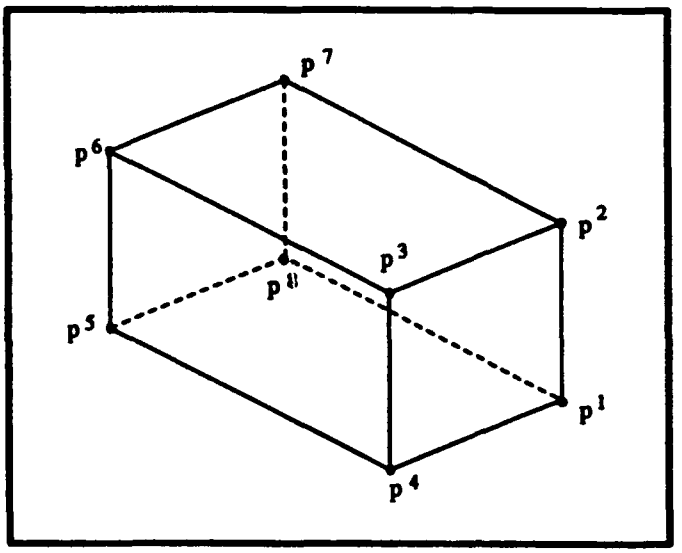

Figure 4.1

Object used in simulation studies. Feature points of interest are labelled p^3, p^5, p^6, and p^7.

object in the image when the object is closest to the cameras. Measurement models in the Kalman filters assume Gaussian statistics with standard deviation 0.02 units in both horizontal and vertical components. As a result, the measurement models are not exactly matched to the true statistics of measurement error, as would probably be the case in any implementation.

Section 4.1 provides a brief investigation of Cramer Rao bounds, first for the resetting requirement of the angle-axis filter, and second to compare optimal performance for the three filters. Section 4.2 provides an overview of simulations and results which are presented in detail in Section 4.3. A closing summary and discussion is provided in Section 4.4.

4.1 Comparison of Cramer Rao Bounds

As discussed in Section 2.2, the extended Kalman filter is generally a nonlinear, suboptimal state estimation technique. Performance assessment of such techniques often employ the Cramer-Rao inequality [79] to derive covariance lower bounds. Appendix C reviews the significance and derivation of Cramer-Rao lower bounds (CRLB) for this estimation problem. Generally, the error covariance matrix $\mathbf{S}(N)$ of an *unbiased* estimator $\hat{\mathbf{x}}(N)$ of the state $\mathbf{x}(N)$, where the estimate is based on the prior information in $\mathbf{x}_0$ and the observations $\mathbf{z}(0),\ldots,\mathbf{z}(N)$, is lower-bounded by the inverse of Fisher's information matrix [80, pp. 91-93]. This result, in theory, provides estimation error variance lower bounds for elements of $\hat{\mathbf{x}}(N)$. It should be noted,

however, that these bounds assume unbiased state estimation; a state dependent bias can result in higher or lower bounds [78] (see Appendix C). As a result, most analysis [28]-[31], [61], [62], proceed under the assumption of unbiased estimation resulting in approximate covariance bounds.

A further assumption in deriving Cramer Rao bounds is that the dynamic model is noise-free. This means that the dynamic model from which the filter is constructed is assumed to exactly match the deterministic temporal behavior of observed processes. For the purposes of satisfying this assumption during investigations of this Section, modeling parameters α_t, α_r, σ_t, and σ_r are set to zero, thereby eliminating process noise and reducing the dynamic models to truncated Taylor series expansions, the initialization parameter vector $\mathbf{d}_O$ is set to zero, which means that the origin of the object centred reference frame lies at a feature point, and true object motion is defined such that it can be exactly modelled by any of the three parameterizations of rotational motion. These restrictions obviously place some limitations on the overall scope of evaluations with Cramer Rao bounds, however, comparisons of optimal performance possible under these ideal circumstances provides a general appreciation for accuracy that can be expected for more complex motion investigated in Sections 4.2 and 4.3.

It was noted in Section 3.5.1 that the orientation vector of the angle-axis filter must be reset, first to maintain a finite magnitude (rotation angle) when the object continually rotates in one direction, and second to yield decreasing Cramer Rao bounds. Figure 4.2 shows Cramer Rao bounds in the absence of prior information ($\mathbf{P}_0^{-1} = \mathbf{0}$) for the rotational state vector $\mathbf{x}_r$ defined in (3.41) for purely rotational motion with $N_\xi = 1$, i.e. constant first time derivative of ξ. Here, object structure is assumed to be known, and is removed from the state vector. The origin of the object-centred frame, which lies at the block centroid, remains fixed at $[30, 0, 0]_E^T$ units, with rotational motion about an axis parallel to $\mathbf{z}_E$ defined by $\xi(0) = [0, 0, 0]_E^T$ and $\dot{\xi}(0) = [0, 0, 1]_E^T$ radians/s. Note that in this case the reset is exact since the axis of the rotation matrix, $\bar{\xi}$, is constant.

The top two graphs of Figure 4.2 are shown without applying the reset of (3.45) while the bottom two graphs are shown with resetting. Without the reset, and for this trajectory, the CRLB of the x- and y-components of ξ increase after 2 seconds. The CRLB of the x- and y-components of $\dot{\xi}$ are approximately an order of magnitude larger than the z-component CRLB at the end of the trajectory. Because, for this trajectory, $\bar{\xi}_1 = \bar{\xi}_2 = 0$, the second and fourth terms in expression (D.10) for $\partial \mathbf{I}_E^O / \partial \xi_j$ with $j = 1, 2$ are zero. As $\gamma = \|\xi\|$ increases, $\partial \mathbf{I}_E^O / \partial \xi_j \rightarrow \mathbf{0}$ for $j = 1, 2$ which implies that the rotation matrix becomes insensitive to changes in $\bar{\xi}_1$ and $\bar{\xi}_2$ which, in turn, leads to increasing CRLB's. The bottom two graphs of Figure 4.2 show CRLB's for ξ and $\dot{\xi}$ with resetting applied as shown in (3.45). In this case, the bounds exhibit a decreasing trend over the trajectory with final values after 20 seconds of approximately 0.005 radians for ξ and 0.0005 radians/s for $\dot{\xi}$.

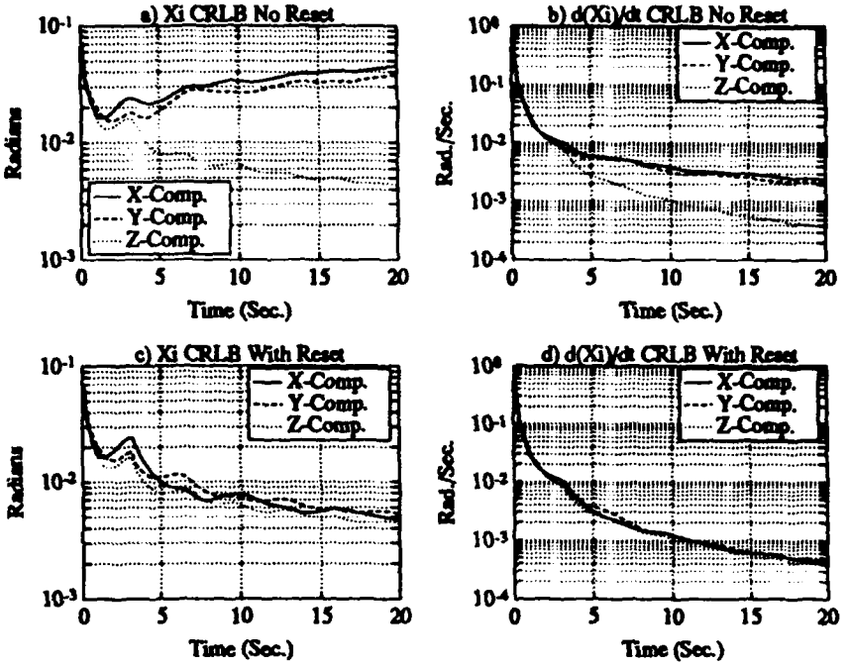

Figure 4.2

CRLB's of orientation vector a) without and b) with resetting

Figures 4.3 through 4.6 show Cramer Rao bounds for the three filters in the absence of prior information ($\mathbf{P}_0^{-1} = \mathbf{0}$) during a simple trajectory with constant axis of rotation. Measurements of positions of feature points p^3, p^5, p^6, and p^7 (see Figure 4.1) are available throughout the trajectory. The initialization parameter vector $\mathbf{d}_O$ is set to zero, point p^6 is chosen to lie at the origin of the object-centred reference frame, point p^7 lies on the x_O-axis, and point p^3 lies in the x_O-y_O plane. All three coordinates of p^5 are unknown. True motion of the object is defined by constant acceleration translational motion of point p^6,

$$\begin{aligned} \mathbf{p}(0) &= [30; -20; 0]_E^T \text{ units}, \\ \dot{\mathbf{p}}(0) &= [2; 2; -2]_E^T \text{ units/s, and} \\ \ddot{\mathbf{p}} &= [-0.4; 0.4; 0.4]_E^T \text{ units/s}^2, \end{aligned} \tag{4.1}$$

and rotational motion about the $\mathbf{y}_O$-axis with constant angular acceleration,

Angle-Axis -

$$\xi(0) = \begin{bmatrix} 0 \\ -\pi/4 \\ 0 \end{bmatrix}, \quad \dot{\xi}(0) = \begin{bmatrix} 0 \\ 0.8\pi \\ 0 \end{bmatrix}, \quad \ddot{\xi} = \begin{bmatrix} 0 \\ -0.16\pi \\ 0 \end{bmatrix}, \tag{4.2}$$

Quaternion -

$$\mathbf{q}(0) = \begin{bmatrix} 0 \\ \sqrt{2} \\ 0 \\ \sqrt{2} \end{bmatrix}, \quad \mathbf{w}(0) = \begin{bmatrix} 0 \\ 0.8\pi \\ 0 \end{bmatrix}, \quad \dot{\mathbf{w}} = \begin{bmatrix} 0 \\ -0.16\pi \\ 0 \end{bmatrix}, \tag{4.3}$$

Roll-Pitch-Yaw $\quad (\zeta = [\phi, \theta, \psi]^T)$

$$\zeta(0) = \begin{bmatrix} 0 \\ -\pi/4 \\ 0 \end{bmatrix}, \quad \dot{\zeta}(0) = \begin{bmatrix} 0 \\ 0.8\pi \\ 0 \end{bmatrix}, \quad \ddot{\zeta} = \begin{bmatrix} 0 \\ -0.16\pi \\ 0 \end{bmatrix}. \tag{4.4}$$

The angle-axis filter is reset twice during this trajectory at about 2.1 and 8.2 seconds, and pitch in the roll-pitch-yaw filter passes through the singular point of $\theta = \pi/2$ at about 1 and 9 seconds. True structure parameters in this case (measured in units) are

$$\mathbf{r}_O^5 = \begin{bmatrix} 0 \\ 0 \\ -10 \end{bmatrix}, \quad \mathbf{r}_O^3 = \begin{bmatrix} 0 \\ 20 \\ 0^* \end{bmatrix}, \quad \mathbf{r}_O^3 = \begin{bmatrix} 10 \\ 0^* \\ 0^* \end{bmatrix}, \tag{4.5}$$

where $(\cdot)^*$ indicates a known value.

Cramer Rao bounds for translational motion estimation, Figure 4.3 and top of Figure 4.4, demonstrate, as might be expected due to identical models for translational motion, that bounds on all three filters are very close for this trajectory, especially for the angle-axis and roll-pitch-yaw filters. Bounds for the quaternion filter lie slightly above those of the other two parameterizations. Bounds on orientation parameters, shown in the lower part of Figure 4.4, exhibit oscillatory behavior, but a decreasing trend over the trajectory. These oscillations appear to contradict the notion that as more measurements become available, more information should be accumulated about the observed process, which means that one might expect the bounds to decrease monotonically. Appendix C clarifies the meaning of these bounds in the present context and demonstrates why they need not decrease monotonically.

Figure 4.5 compares covariance bounds for estimation of angular velocity and acceleration. Again very similar behavior is demonstrated for the angle-axis and roll-pitch-yaw filters, whereas bounds for the quaternion filter are slightly higher. Figure 4.6 shows the bounds for estimation of the six structural parameters. Again, almost identical behavior is shown for the angle-axis and roll-pitch-yaw filters. However, significantly higher bounds are present for the quaternion filter in the estimation of the z_O-coordinate of p^5, the y_O-coordinate of p^3, and the x_O-coordinate of p^7. Note that these are precisely the non-zero parameters in the structure set shown in (4.5). This investigation was repeated for different combinations of structural parameters and always yielded similar results in that quaternion filter bounds for non-zero structural parameters were significantly above those for the other two filters.

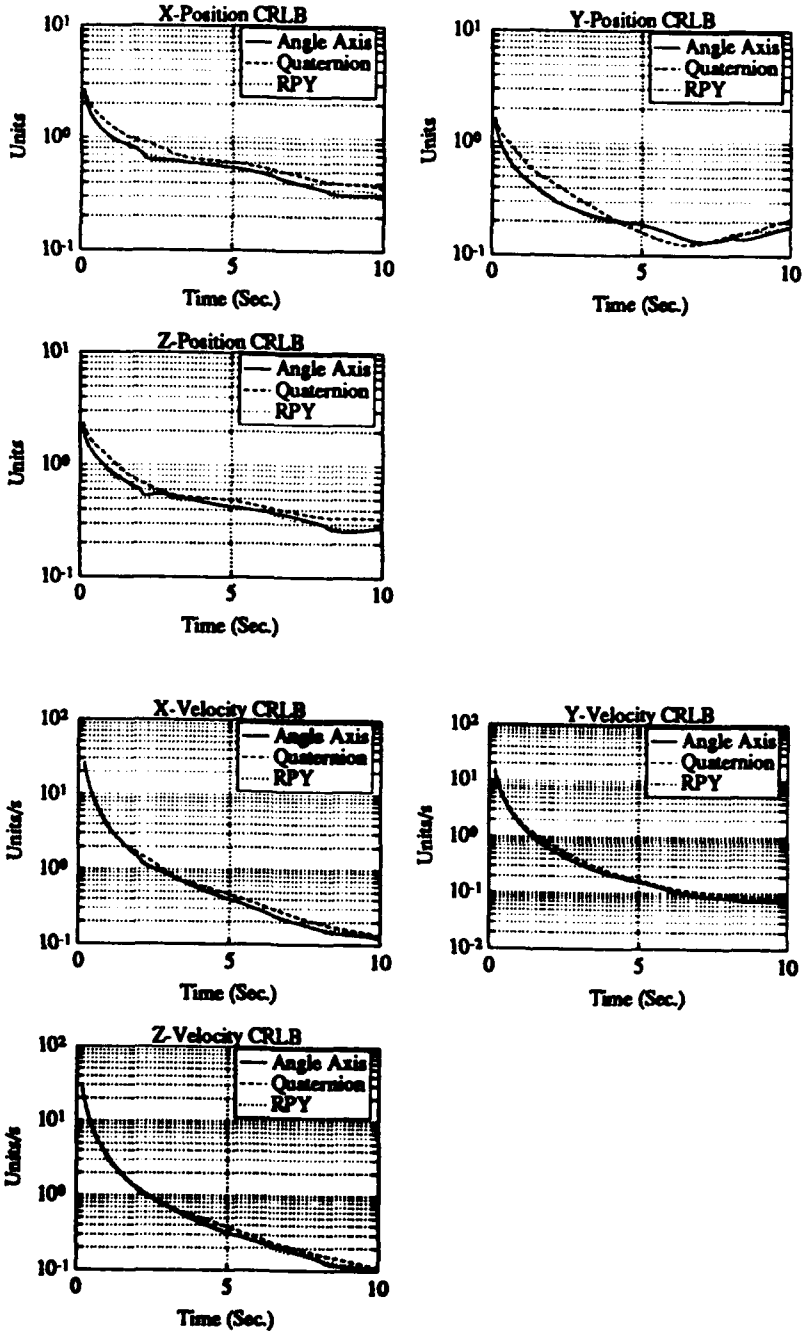

Figure 4.3

Cramer Rao bound comparison of position and velocity.

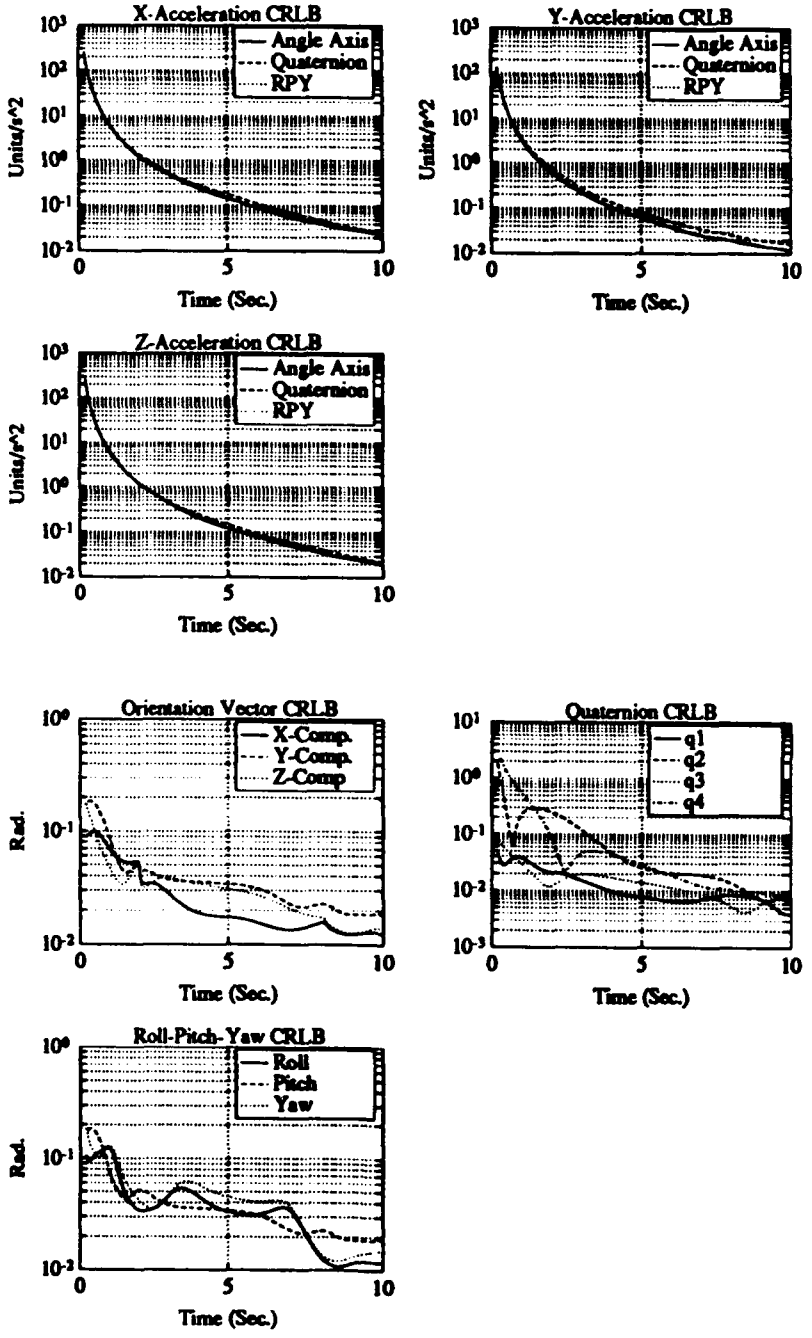

Figure 4.4

Cramer Rao bound comparison of acceleration and orientation.

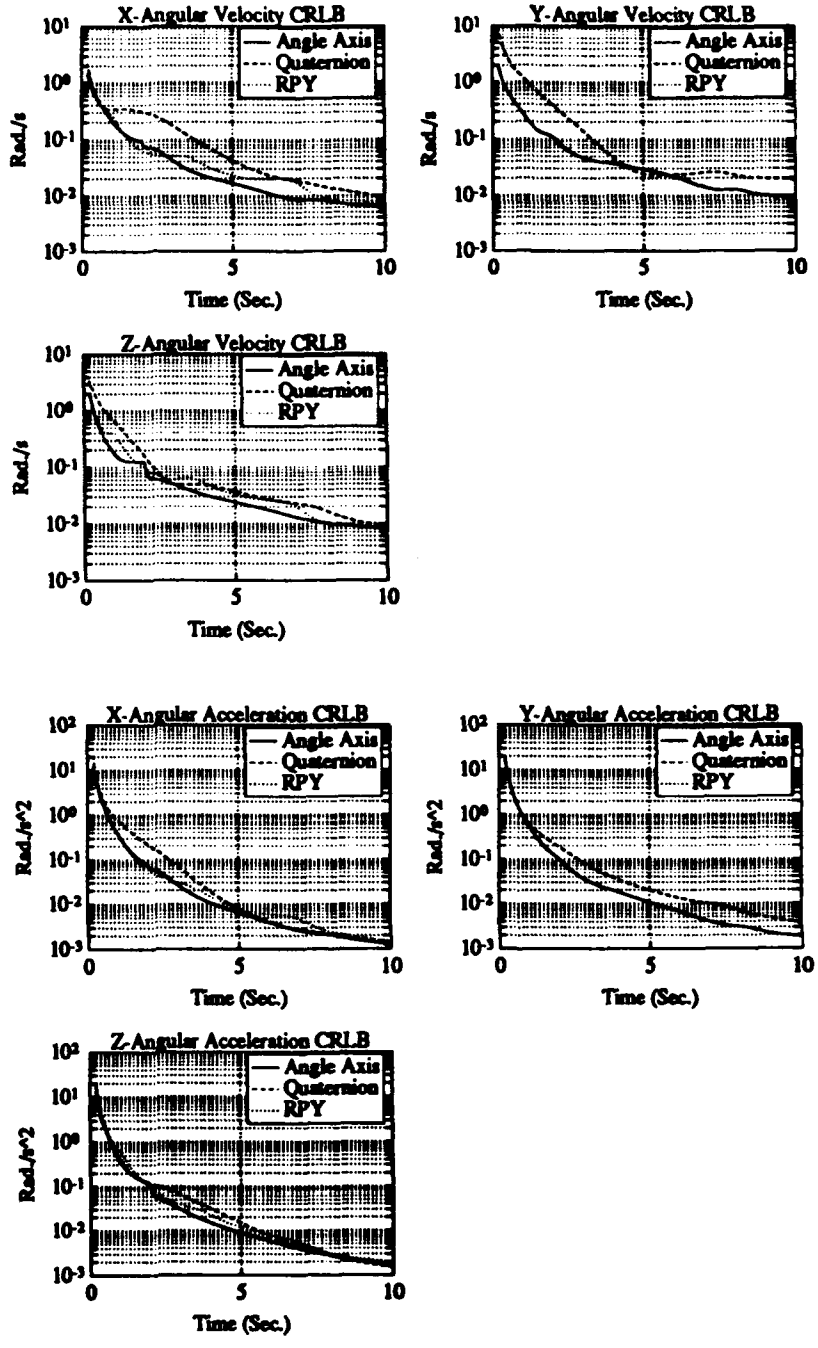

Figure 4.5

Cramer Rao bound comparison of rotational motion.

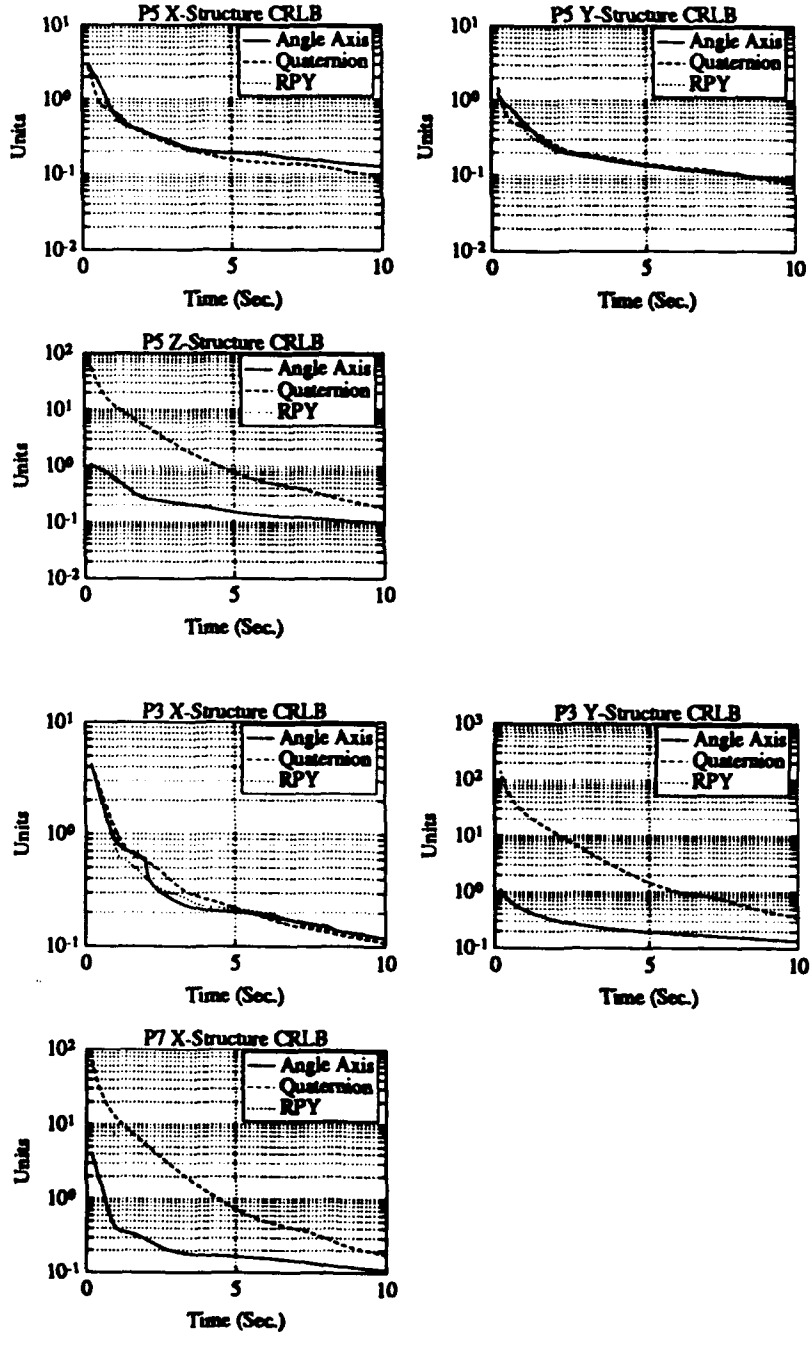

Figure 4.6

Cramer Rao bound comparison of structure estimation.

4.2 Overview of Simulations and Results

Simulation studies have been conducted over a wide range of trajectories, beginning with known structure and simple motion, and progressing to completely unknown structure with more complex motion. Various sets of parameters were employed for dynamical models in the filters during these simulations. A simple process for selecting physically justifiable parameters led to the best results.

Because measurements are related directly only to object position and orientation, estimation of high-order time derivatives often demonstrated serious lag and overshoot which can lead to filter instability. Consequently, motion only up to the second time derivative is included in dynamic models by setting $N_t = 2$, $N_\xi = 2$, $N_w = 1$, and $N_\zeta = 2$. Following Singer's approach [71] in Equation (3.33) of Section 3.4, the manoeuvre correlation time constant is estimated for the expected class of objects and motion, extreme values are estimated for the manoeuvre variables of translational and rotational motion, and probabilities are assigned to the occurrence of maximum, minimum and zero values of the manoeuvre variable. Results shown in Section 4.3 were generated assuming a manoeuvre correlation time constant of 2 seconds which gives $\alpha_t = \alpha_r = 0.5\text{s}^{-1}$. Extreme values of acceleration for translational motion is estimated as $M_{t_{\max}} = 12$ units/s^2 which occur with probability $P_{t_{\max}} = 0.1$, and zero translational acceleration occurs with probability $P_{t_0} = 0.1$. These values in Equation (3.33) give

$$\sigma_t = \left\{ \frac{M_{t_{\max}}^2}{3} [1 + 4P_{t_{\max}} - P_{t_0}] \right\}^{\frac{1}{2}} \approx 8. \tag{4.6}$$

The continuous-time white noise process driving the translational motion model then has covariance $\mathbf{Q}_t = 2\alpha_t\sigma_t^2\mathbf{I}_3$, and the approximation of (3.32) is employed to compute the covariance matrix of the corresponding discrete-time process noise sequence. Corresponding parameters for rotational motion are $M_{r_{\max}} = 6\text{rad/s}^2$, $P_{r_{\max}} = 0.1$, and $P_{r_0} = 0.1$, which gives $\sigma_r \approx 4$ and $\mathbf{Q}_r = 2\alpha_r\sigma_r^2\mathbf{I}_3$. Integration of the quaternion differential equation with $N_\pi = 5$ was found to give sufficient accuracy.

A subset of results from three trajectories are presented in Section 4.3.

Case 1 With model parameters selected as described above, the Extended Kalman filter keeps the gains relatively large in order to follow manoeuvres as they occur. Consequently, although performance for manoeuvring trajectories is good, performance for nonmanoeuvring trajectories may be degraded from that of simpler filters [71]. This Case considers the nonmanoeuvring trajectory employed in the comparison of Cramer Rao bounds in Section 4.1, and shows results of filters exactly "matched" to the trajectory (no process noise) and results from "unmatched" filters formed from the parameter selection above. Results demonstrate that although performance of the unmatched filters are slightly degraded in the estimation of translational and rotational time derivatives, performance

in the estimation of position, orientation, and structure are very similar for both the matched and unmatched filters.

Case 2 True object motion is defined with constant acceleration of the block centroid, and rotational motion computed from an angle-axis formulation which cannot be exactly matched by any of the three parameterizations of rotational motion employed in the filters. The origin of the object-centred reference frame does not, due to the initialization algorithm described in Section 3.6, lie at the block centroid, and hence translational motion models in each filter are not matched to the trajectory. The angle-axis filter gives slightly better overall performance for this trajectory than the quaternion and roll-pitch-yaw filters; the roll-pitch-yaw filter demonstrates poorer performance in estimation of angular velocity, and the quaternion filter exhibits poorer performance in the estimation of position and some structural components.

Case 3 The trajectory of Case 2 was based on an angle-axis formulation in which case one might expect the angle-axis and quaternion filters to more accurately match the trajectory than the roll-pitch-yaw filter. The trajectory of Case 3 defines true object motion with constant acceleration of the block centroid, as in Case 2, but rotational motion is computed from a sinusoidal roll-pitch-yaw formulation. As in Case 2, neither translational nor rotational dynamic models are matched to the trajectory. The angle-axis and roll-pitch-yaw filters show almost identical performance in this case. Again the quaternion filter exhibits poorer performance in the estimation of position and some structural components.

Results from these three Cases are representative of all results obtained from various trajectories investigated during simulation studies. In general, all three filters demonstrated satisfactory tracking of manoeuvring objects with the angle-axis filter giving slightly better performance than the quaternion and roll-pitch-yaw filters. The quaternion filter, apart from its higher computational requirements, appears to give poorer (slower) performance in tracking position and estimation of structure. The roll-pitch-yaw filter, which is the simplest from a computational point of view, sometimes showed significant error in angular velocity estimates. The angle-axis filter is only slightly more computationally demanding than the roll-pitch-yaw filter. The following Section presents detailed graphical results for the above three cases, and provides further discussions on relative performance of the three filters.

4.3 Detailed Simulation Results

This Section presents partial sets of simulation results for the three Cases outlined in Section 4.2. The vision system and object geometry, and measurement noise levels are as described in the Introduction to this Chapter. With four feature points of

interest, the angle-axis and roll-pitch-yaw filters have 24 states, while the quaternion filter has 25 states. Presentation of results for each Case includes a discussion of the true object motion, initialization parameters, true structure (which depends on the initialization parameter vector $\mathbf{d}_O$), and the resultant image sequences. Graphical presentation may include results of single-run samples and/or 60-run Monte Carlo simulations. In order to duplicate true motion and structure over Monte Carlo simulations, initialization during the first run selects $\mathbf{d}_O$ (the same for all three filters), and subsequent runs employ the same $\mathbf{d}_O$ as the first, but with other parameters being initialized as outlined in Section 3.6.

The angle-axis and roll-pitch-yaw formulations provide intuitive parameterizations for the three-dimensional rotational group. However, to say that an estimated quaternion $\hat{\mathbf{q}}$ has an error $\tilde{\mathbf{q}}$, what does this mean in terms of the error in the corresponding basis transformation $\tilde{\mathbf{I}}_E^O \triangleq \mathbf{I}_E^O - \hat{\mathbf{I}}_E^O$? A convenient measure of estimation accuracy at the kth time step is given by the scaled Frobenius norm

$$C(k) \triangleq \frac{1}{2\sqrt{3}} \left(\mathrm{Tr}\{\tilde{\mathbf{I}}_E^O(k)(\tilde{\mathbf{I}}_E^O)^T(k)\} \right)^{\frac{1}{2}}. \tag{4.7}$$

The Frobenius matrix norm without the normalization factor $1/(2\sqrt{3})$ has been employed in investigations such as [65]. Because $\mathbf{I}_E^O(k)$ and $\hat{\mathbf{I}}_E^O(k)$ are both orthogonal for all three parameterizations, the scale factor employed in (4.7) implies that the smallest interval containing $C(k)$ is $[0, 1]$.

Angular velocity is available directly in the quaternion filter, and is computed from state estimates using Equations (3.40) and (3.85) for the angle-axis and roll-pitch-yaw filters, respectively. Errors in angular velocity estimation are shown as sample or root-mean-square values of the Euclidean norm of the difference between true and estimated angular velocity vectors. Sample results show true and estimated values of position and velocity, whereas Monte Carlo simulation results show root-mean-square values of the Euclidean norm of the difference between true and estimated position and velocity vectors, respectively. Results for structure estimation show estimation errors for sample runs and root-mean-square values of the Euclidean norm of the difference between true and estimated structure vectors for Monte Carlo simulations.

Case 1: This Case considers the nonmanoeuvring trajectory employed in the comparison of Cramer Rao bounds in Section 4.1, and shows results of filters exactly "matched" to the trajectory (no process noise) and results from "unmatched" filters formed from parameter selection as shown in Table I.

The matched filters set $\alpha_t = \alpha_r = 0\ \mathrm{s}^{-1}$, $\sigma_t = 0$ units/s^2, and $\sigma_r = 0$ rad/s^2, while the unmatched filters employ $\alpha_t = \alpha_r = 0.5\ \mathrm{s}^{-1}$, $\sigma_t = 8$ units/s^2, and $\sigma_r = 4$ rad/s^2. The initialization parameter vector $\mathbf{d}_O$ is set to zero for this investigation. True translational motion of the origin of the object-centred frame, which lies at feature point p^6, true rotational motion about p^6, and true structure parameters are given in Table I along with a sample initialization denoted by $\hat{\mathbf{x}}_{s,t,r}(0)$. Note that structure

Table I

Simulation parameters for Case 1. Table entries replaced by 'NC' indicate that quantities have not been computed and are not used in initialization.

Structure: $N_f = 4$, $\sigma(\tilde{\mathbf{x}}_s) = 5.0$, $\mathbf{d}_0 = [0,0,0]^T$			
	$\mathbf{r}_O^5$	$\mathbf{r}_O^3$	$\mathbf{r}_O^7$
$\mathbf{x}_s$	$[0,0,10]$	$[0,20,d_{O,3}]$	$[10,d_{O,2},d_{O,3}]$
$\hat{\mathbf{x}}_s(0)$	$[-4.54,-0.30,9.15]$	$[-11.56,18.87,d_{O,3}]$	$[17.96,d_{O,2},d_{O,3}]$

Translation: $N_t = 2$, $\alpha_t = 0$ and 0.5, $\sigma_t = 0$ and 8			
	$\mathbf{p}$	$\dot{\mathbf{p}}$	$\tilde{\mathbf{p}}$
$\mathbf{x}_t(0)$	$[30,-20,0]$	$[2,2,-2]$	$[-0.4,0.4,0.4]$
$\hat{\mathbf{x}}_t(0)$	$[31.84,-20.72,2.23]$	$[0,0,0]$	$[0,0,0]$
$\sigma(\tilde{\mathbf{x}}_t)$	5.0	5.0	5.0

Rotation: Angle-Axis, $N_\xi = 2$, $\alpha_r = 0$ and 0.5, $\sigma_r = 0$ and 4			
	ξ	$\dot{\xi}$	$\tilde{\xi}$
$\mathbf{x}_r(0)$	$[0,-0.7854,0]$	NC	NC
$\hat{\mathbf{x}}_r(0)$	$[0.0052,-0.72,-0.2]$	$[0,0,0]$	$[0,0,0]$
$\sigma(\tilde{\mathbf{x}}_r)$	0.5	2.0	2.0

Rotation: Quaternion, $N_w = 1$, $N_x = 5$, $\alpha_r = 0$ and 0.5, $\sigma_r = 0$ and 4			
	$\mathbf{q}$	$\mathbf{w}$	$\dot{\mathbf{w}}$
$\mathbf{x}_r(0)$	$[0,0.38,0,0.92]$	NC	NC
$\hat{\mathbf{x}}_r(0)$	$[0.0025,0.35,0.098,0.93]$	$[0,0,0]$	$[0,0,0]$
$\sigma(\tilde{\mathbf{x}}_r)$	0.5	2.0	2.0

Rotation: Roll-Pitch-Yaw, $N_\zeta = 2$, $\alpha_r = 0$ and 0.5, $\sigma_r = 0$ and 4			
$\zeta = [\phi,\theta,\psi]$	ζ	$\dot{\zeta}$	$\tilde{\zeta}$
$\mathbf{x}_r(0)$	$[0,-0.7854,0]$	NC	NC
$\hat{\mathbf{x}}_r(0)$	$[0.084,-0.72,-0.24]$	$[0,0,0]$	$[0,0,0]$
$\sigma(\tilde{\mathbf{x}}_r)$	0.5	2.0	2.0

initialization for the sample shown has significant errors, particularly in $\mathbf{r}^7_{O,1}$. The initial filter covariance matrices are taken as diagonal with the square-roots of diagonal entries corresponding to standard deviations, as indicated in Table I, of 5 units for structure states, 5 units for position and its time derivatives, 0.5 for orientation, and 2.0 Rad/s for time derivatives of orientation parameters.

True and noisy feature point trajectories in the image plane of the left camera are shown in Figure 4.7. In the image, the top of the object first rotates towards the observer, slows down, stops, and then reverses its direction of rotation to return to its original orientation. The axis of rotation is constant for this trajectory. Measurement noise, illustrated in the lower half of Figure 4.7 is moderately high, but free of significant outliers due to the assumption of uniformly distributed errors. In practice, the frequency of occurrence of outlier measurements for object tracking filters could be reduced through a hierarchical approach similar to that outlined in Section 1.2, or rejected with chi-squared tests as employed by Ayache and Faugeras [40].

Figures 4.8 and 4.9 show 60-run Monte Carlo simulation results for matched and unmatched filters, respectively. In the absence of process noise, Figure 4.8, the Kalman filters eventually neglect further measurements as information is accumulated. This is demonstrated by the sample trace (lower right graph of Figure 4.8) of the maximum principal value (singular value) of the Kalman gain matrix over time. For this particular trajectory and measurement noise, the filter gains appear to prematurely reject further measurements. This is demonstrated by the increasing root-mean-square errors in position estimation, with the angle-axis filter showing the poorest performance in the second half of the trajectory. Cramer Rao bounds for position estimation at the end of the trajectory, from Figure 4.3, are approximately 0.3 units in each component, or very roughly $0.3\sqrt{3} \approx 0.5$ units for the Euclidean norm of position errors. This suggests reasonably good performance of the quaternion and roll-pitch-yaw filters in position estimation. The quaternion filter shows slightly better performance in translational velocity and orientation estimation, while all three filters yield similar performance, in comparison to Cramer Rao bounds, for angular velocity and structure estimation.

Figure 4.9 shows results for the three filters with parameters selected to track manoeuvring objects. All three filters demonstrate similar performance during this simulation. In this case the maximum principal gains are maintained at approximately 95 to 100 after initial transients. In comparison to performance of the matched filters of Figure 4.8, position orientation and structure estimation is very similar, but translational and rotational velocity estimation is degraded. Generally, higher filter gains were found to increase root-mean-square errors, but decrease mean errors (bias) in velocity and higher time derivatives. These results demonstrate the conflicting requirements of keeping the gains small enough to suppress measurement noise while at the same time keeping the gains high enough to track manoeuvring objects or to maintain robustness in the presence of unmodelled dynamics.

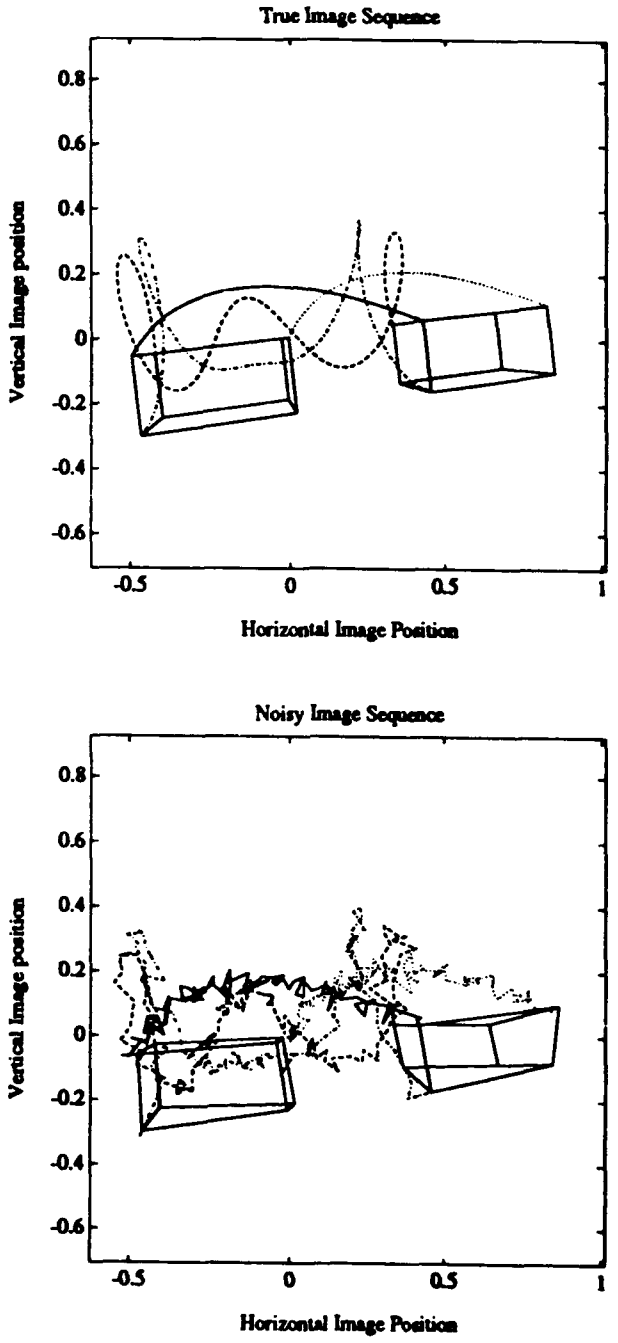

Figure 4.7

True and sample noisy image feature point trajectories in the left camera without occlusion for Case 1.

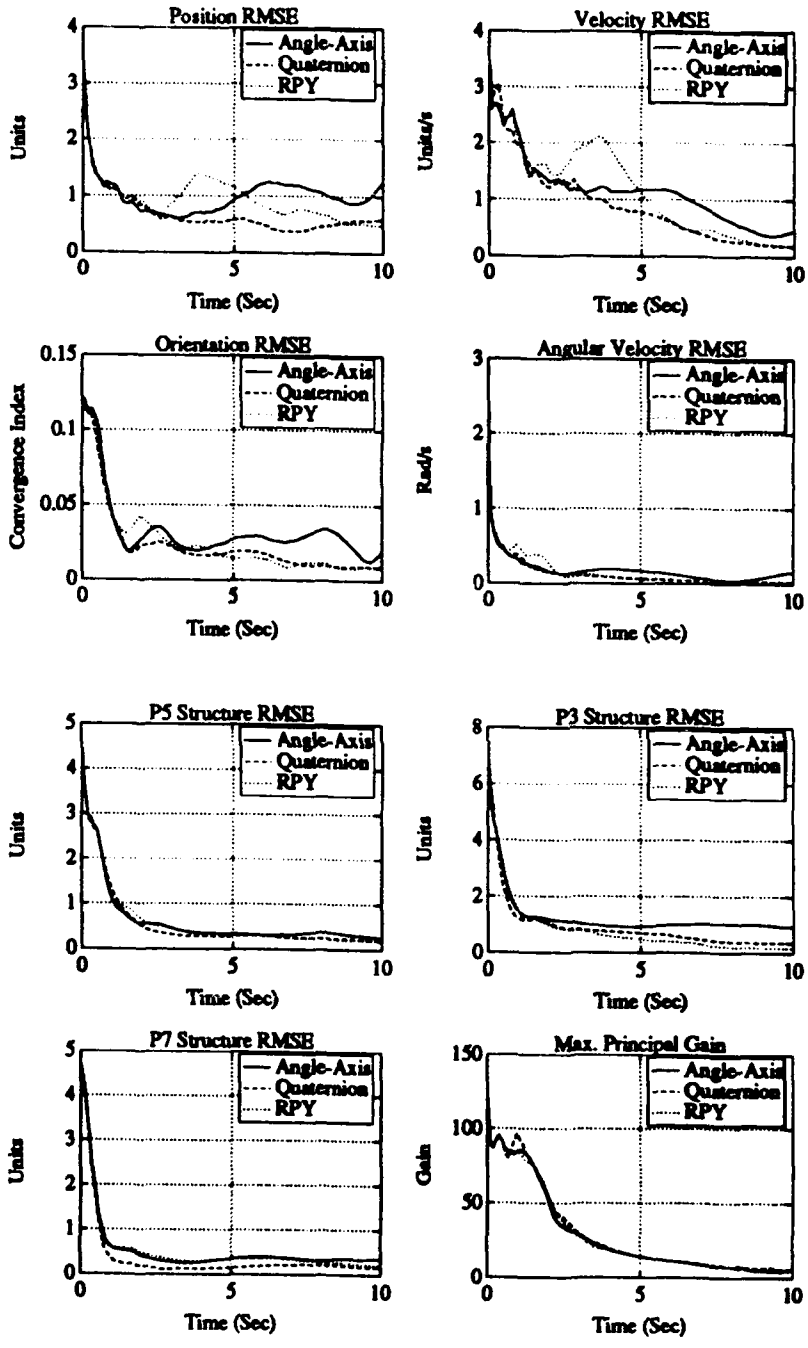

Figure 4.8

Monte Carlo simulation results for Case 1 without process noise.

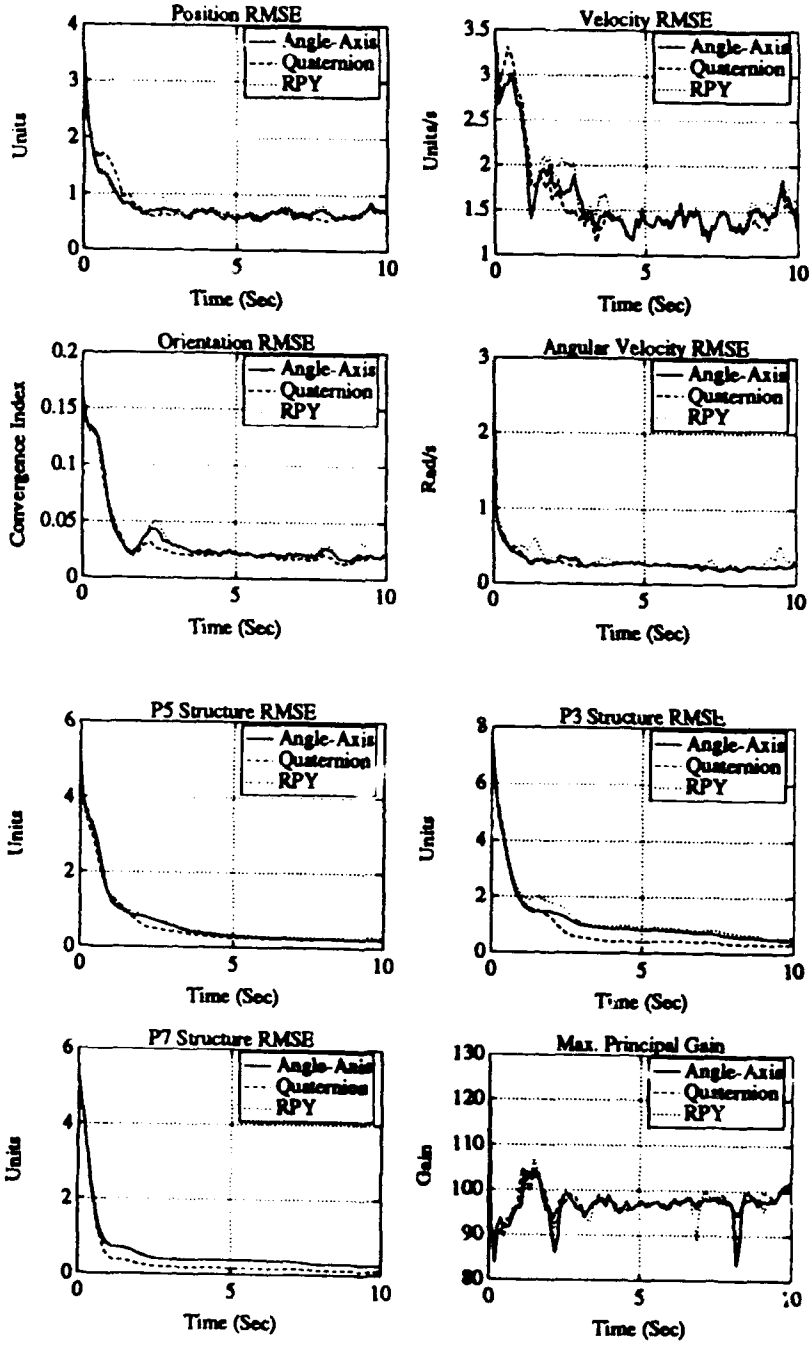

Figure 4.9

Monte Carlo simulation results for Case 1 with process noise.

Case 2: In this case, true object motion is defined with constant acceleration of the block centroid, and rotational motion computed from an angle-axis formulation. For the purposed of describing the true motion of the block, define F_B as the block-centred reference frame used in the imagery generation program. Note that F_B and F_O are generally distinct, but for the purpose of simulation differ only by a translation (so that $\mathbf{I}_B^O \equiv \mathbf{I}_3$). The origin, O_B, of F_B lies at the block centroid, and the coordinate planes of F_B are parallel to faces of the block. Translational motion of O_B was defined in the imagery generation program with

$$\begin{aligned} [\mathbf{T}_B^E]_E(0) &= [30, -20, 0]_E^T \text{ units}, \\ [\dot{\mathbf{T}}_B^E]_E(0) &= [2, 2, -2]_E^T \text{ units/s, and} \\ [\ddot{\mathbf{T}}_B^E]_E &= [-0.4, 0.4, 0.4]_E^T \text{ units/s}^2. \end{aligned} \tag{4.8}$$

Recall that the initialization routine computes $\mathbf{d}_O$, which defines the position of the origin of the object centred frame, based on observations during the first measurement event. The origin of the object-centred frame employed in the filters, therefore, will not generally coincide with the centroid of the block and, due to rotational motion, will **not** have the constant acceleration motion defined in (4.8). Specifically, the filters are required to estimate translational motion of F_O whose position and velocity, for example, are given by

$$\mathbf{p}(t) \triangleq [\mathbf{T}_O^E]_E(t) = [\mathbf{T}_B^E]_E(t) + \mathbf{I}_E^O \mathbf{I}_O^B \left\{ \mathbf{r}_B^{i_0} - [\mathbf{d}_O]_B \right\}, \tag{4.9}$$

$$\dot{\mathbf{p}}(t) = [\dot{\mathbf{T}}_B^E]_E + [\mathbf{w}_E^O]_E^{\natural} \mathbf{I}_E^O \mathbf{I}_O^B \left\{ \mathbf{r}_B^{i_0} - [\mathbf{d}_O]_B \right\}, \tag{4.10}$$

where the velocity equation uses (3.35) and the fact that $\mathbf{I}_B^O$ is constant. Higher-order time derivatives of true translational motion are computed in a straight-forward manner by differentiating (4.10).

Rotational motion of the block-centred reference frame is defined in the imagery generation program with an angle-axis parameterization,

$$\mathbf{I}_E^B(t) = \exp(\eta(t)\bar{\mathbf{u}}^{\natural}(t)). \tag{4.11}$$

The angle of rotation, $\eta(t)$, has constant second time derivative,

$$\eta(0) = 0 \text{ rad}, \quad \dot{\eta}(0) = 0.4\pi \text{ rad/s}, \quad \text{and} \quad \ddot{\eta} = -0.08\pi \text{ rad/s}^2. \tag{4.12}$$

Temporal behavior of $\bar{\mathbf{u}}$, since it is a unit vector, must be rotational with instantaneous angular velocity, say, $\mathbf{w}_u(t)$ in the earth-fixed frame so that

$$\dot{\bar{\mathbf{u}}}(t) = \mathbf{w}_u^{\natural}(t)\bar{\mathbf{u}}, \tag{4.13}$$

and the initial condition is specified as

$$\bar{\mathbf{u}}(0) = \frac{1}{\sqrt{3}}[1,\ 1,\ 1]^T. \tag{4.14}$$

The angular velocity, $\mathbf{w}_u$, of the axis of the basis transformation (this is not the angular velocity of the object) is defined with constant second time derivative,

$$\mathbf{w}_u(0) = \begin{bmatrix} -0.2 \\ -0.2 \\ 0.2 \end{bmatrix} \text{rad}, \quad \dot{\mathbf{w}}_u(0) = \begin{bmatrix} -0.1 \\ -0.1 \\ 0.1 \end{bmatrix} \text{rad/s}, \quad \ddot{\mathbf{w}}_u = \begin{bmatrix} 0.02 \\ 0.02 \\ -0.02 \end{bmatrix} \text{rad/s}^2. \qquad (4.15)$$

Integration of (4.13) is performed numerically using a method similar in form to the quaternion integration scheme described in Section 3.5.2 by defining a partition of each sample period, assuming $\mathbf{w}_u$ is approximately constant over each interval of the partition, and using the closed form expression for the matrix exponential of a 3×3 skew symmetric matrix given in (3.39). As in the quaternion integration scheme, this method ensures that $\bar{\mathbf{u}}(t)$ remains a unit vector.

Table II summarizes simulation parameters for this Case and provides numerical values of a sample initialization. A subsequence of images viewed in the left camera is shown in Figure 4.10. In this Figure, each frame sample number is shown in the lower left: the sequence proceeds across the top row from left to right, back across the second row from right to left, etc. The top of the object first begins to rotate towards the observer as does the right side of the object. The object eventually returns to its original orientation at the end of the trajectory. True and sample noisy feature point trajectories in the image plane of the left camera are shown in Figure 4.11. Note the tight loops which would have to be tracked by temporal correspondence algorithms such as two dimensional feature point trackers in a hierarchical approach similar to that proposed in Section 1.2.

Sample results shown in Figures 4.12 and 4.13 correspond to the initialization listed in Table II. Figure 4.12 shows true and estimated position and velocity. Accurate tracking of position is provided by all three filters, however, velocity estimation indicates slight lag and overshoot, particularly in the z-component. Again the maximum principal gain is maintained in the interval 90 to 100 during the second half of the trajectory. Figure 4.13 shows sample results for orientation, angular velocity, and structure estimation. Similar behavior of the three filters is shown orientation estimation with errors below 5% after initial transients. Both the angle-axis and quaternion filter track angular velocity with errors less than 0.5 rad/s, but the roll-pitch-yaw filter shows a fairly significant spike at approximately 2.5 seconds. All structure estimates show convergence near the end of the trajectory.

Figure 4.14 shows root-mean-square errors from a 60-run Monte Carlo simulation. Again very similar performance is demonstrated by all three filters, except for slightly degraded estimation of position and p^5 structure from the quaternion filter. Note the fairly significant oscillations in translational velocity estimation due to lag and some overshoot in filter estimates. The amplitude of oscillation appears to decrease during the second half of the trajectory. Angular velocity estimates from the roll-pitch-yaw filter shows distinct peaks in root-mean-squared errors at about 2.5, 4.5, and 6 seconds. The true pitch angle (not shown) for this trajectory behaves like a damped

Table II

Simulation parameters for Cases 2 and 3. Table entries replaced by 'NC' indicate that quantities have not been computed and are not used in initialization.

Structure: $N_f = 4$, $\sigma(\tilde{\mathbf{x}}_s) = 5.0$, $\mathbf{d}_0 = [-3.29, -4.74, -3.06]^T$			
	$\mathbf{r}_O^5$	$\mathbf{r}_O^3$	$\mathbf{r}_O^7$
$\mathbf{x}_s(0)$	$[-3.28, -4.74, 6.94]$	$[-3.29, 15.26, d_{O,3}]$	$[6.71, d_{O,2}, d_{O,3}]$
$\hat{\mathbf{x}}_s$	$[-3.12, -4.80, 6.58]$	$[-4.37, 16.05, d_{O,3}]$	$[9.94, d_{O,2}, d_{O,3}]$

Translation: $N_t = 2$, $\alpha_t = 0.5$, $\sigma_t = 8$			
	$\mathbf{p}$	$\dot{\mathbf{p}}$	$\bar{\mathbf{p}}$
$\mathbf{x}_t(0)$	$[28.29, -25.26, -1.94]$	NC	NC
$\hat{\mathbf{x}}_t(0)$	$[28.39, -25.70, -2.05]$	$[0, 0, 0]$	$[0, 0, 0]$
$\sigma(\tilde{\mathbf{x}}_t)$	5.0	5.0	5.0

Rotation: Angle-Axis, $N_\xi = 2$, $\alpha_r = 0.5$, $\sigma_r = 4$			
	ξ	$\dot{\xi}$	$\bar{\xi}$
$\mathbf{x}_r(0)$	$[0, 0, 0]$	NC	NC
$\hat{\mathbf{x}}_r(0)$	$[-0.048, -0.047, -0.040]$	$[0, 0, 0]$	$[0, 0, 0]$
$\sigma(\tilde{\mathbf{x}}_r)$	0.5	2.0	2.0

Rotation: Quaternion, $N_w = 1$, $N_\pi = 5$, $\alpha_r = 0.5$, $\sigma_r = 4$			
	$\mathbf{q}$	$\mathbf{w}$	$\dot{\mathbf{w}}$
$\mathbf{x}_r(0)$	$[0, 0, 0, 1]$	NC	NC
$\hat{\mathbf{x}}_r(0)$	$[0.024, 0.023, 0.02, 0.999]$	$[0, 0, 0]$	$[0, 0, 0]$
$\sigma(\tilde{\mathbf{x}}_r)$	0.5	2.0	2.0

Rotation: Roll-Pitch-Yaw, $N_\zeta = 2$, $\alpha_r = 0.5$, $\sigma_r = 4$			
$\zeta = [\phi, \theta, \psi]$	ζ	$\dot{\zeta}$	$\bar{\zeta}$
$\mathbf{x}_r(0)$	$[0, 0, 0]$	NC	NC
$\hat{\mathbf{x}}_r(0)$	$[-0.047, 0.048, -0.039]$	$[0, 0, 0]$	$[0, 0, 0]$
$\sigma(\tilde{\mathbf{x}}_r)$	0.5	2.0	2.0

sinusoid with peaks just prior to these times, and has a maximum value of about 0.4π. Although estimation of pitch angle was very accurate, estimation of the time derivative of pitch, which is also a damped sinusoid, shows some lag and overshoot much like those occurring in translational velocity estimation. This behavior would contribute significantly to the peaks present in angular velocity estimation error. The angle-axis filter gives marginally better performance than the other two filters for this simulation. Structure root-mean-square estimation errors decrease rapidly during the first 1 or 2 seconds, and appear to have levelled off at about 0.3 to 0.5 units at the end of the trajectory.

As discussed in Section 3.2, there are four possible orientations for the object-centred reference frame with respect to the three special feature points selected at initialization. When structure was initialized with all zeros, for example, convergence of structure estimates was observed in one of the four possible orientations of the object-centred frame, the particular one of the four being dependent on initial filter transients. In simulations with this object and the given measurement noise levels it was found that the simple single-frame initialization of Section 3.6 provided sufficient initial accuracy for structure to converge in the object-centred frame selected at initialization, although obviously this need not always be the case. Figure 4.15 shows the reliable convergence of structure estimates to true structural parameters given in Table II after 10 seconds (100 measurement events) for each of the Monte-Carlo simulation runs. Again all three filters give similar results, with final estimates from the quaternion filter being slightly noisier than those from the angle-axis and roll-pitch-yaw filters.

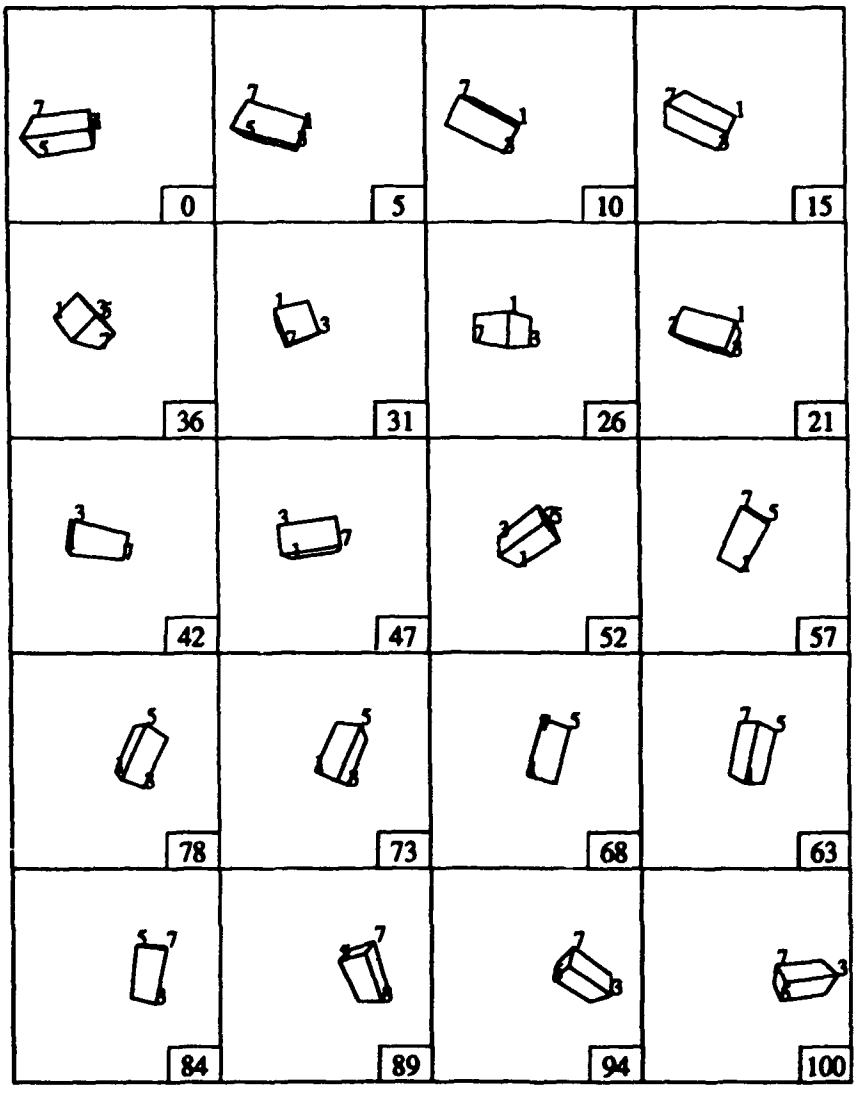

Figure 4.10

Image subsequence from the left camera for Case 2. Image frame numbers are shown in the lower left corner of each image. Four feature points, p^1, p^3, p^5, and p^7, are labelled in each image.

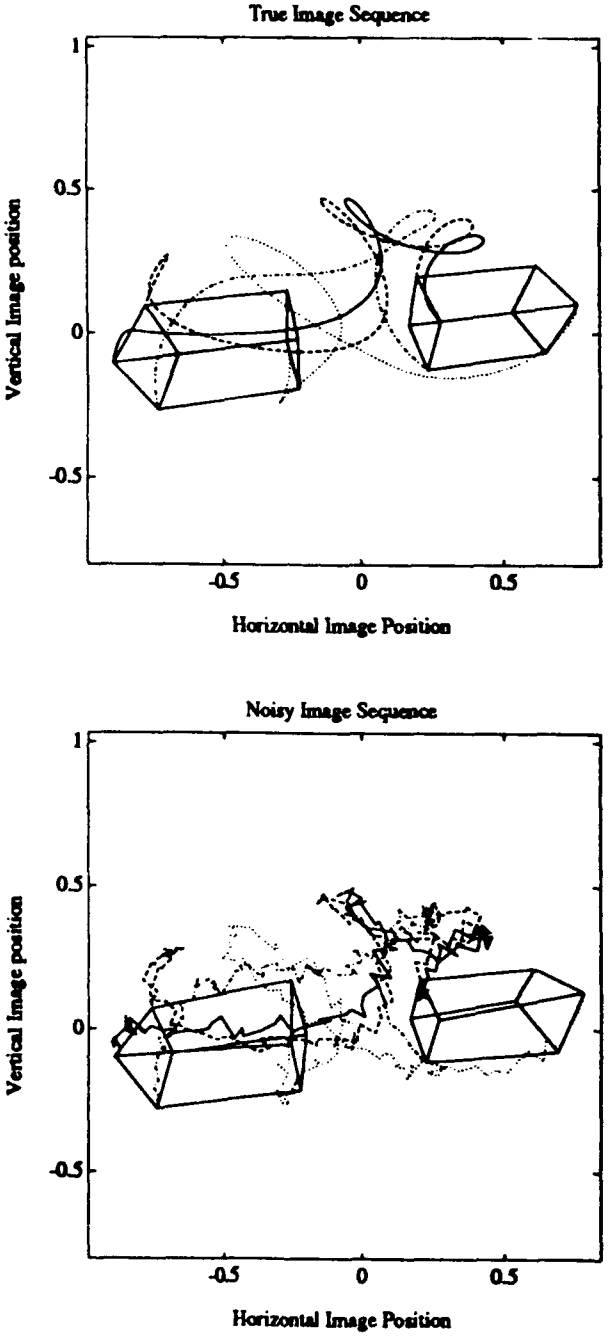

Figure 4.11

True and sample noisy image feature point trajectories in the left camera without occlusion for Case 2.

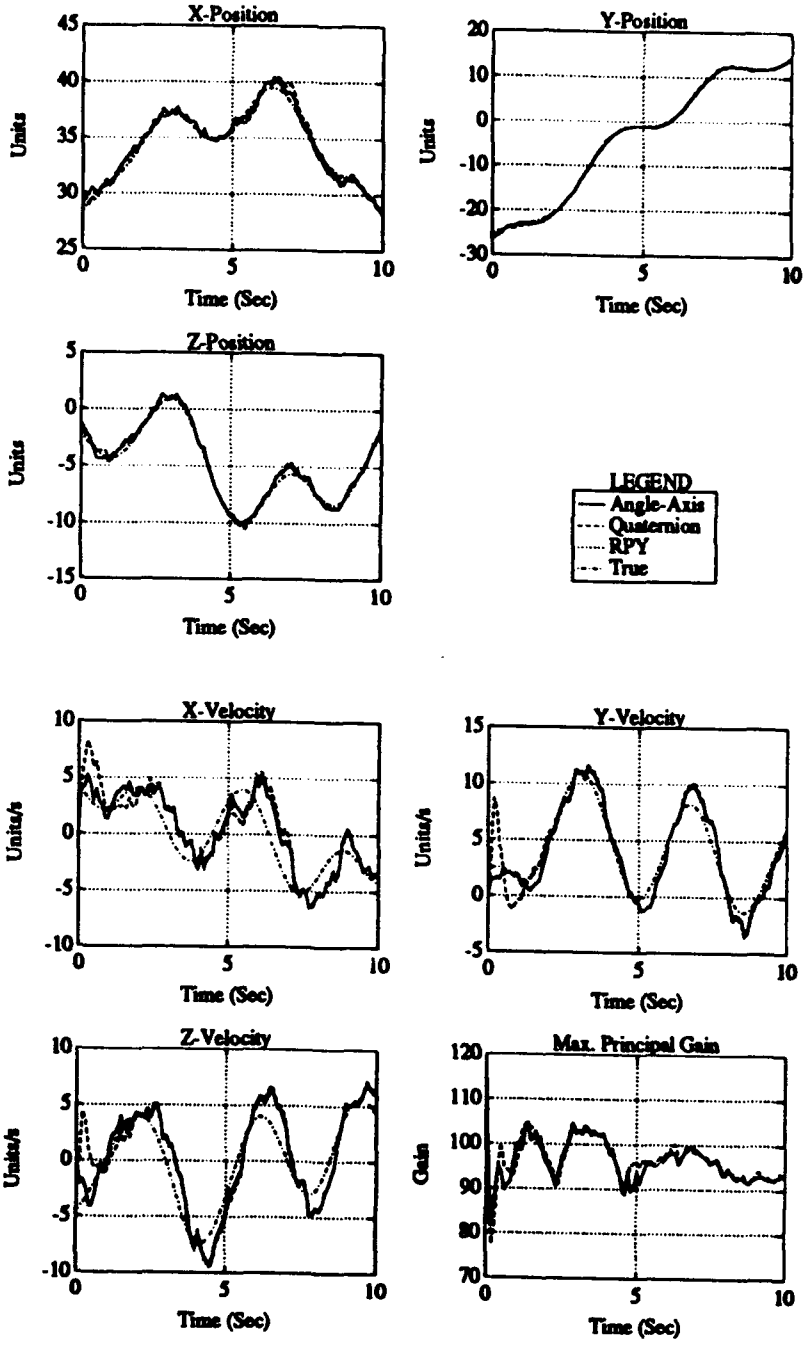

Figure 4.12

Sample results for Case 2, position, velocity and gain strength.

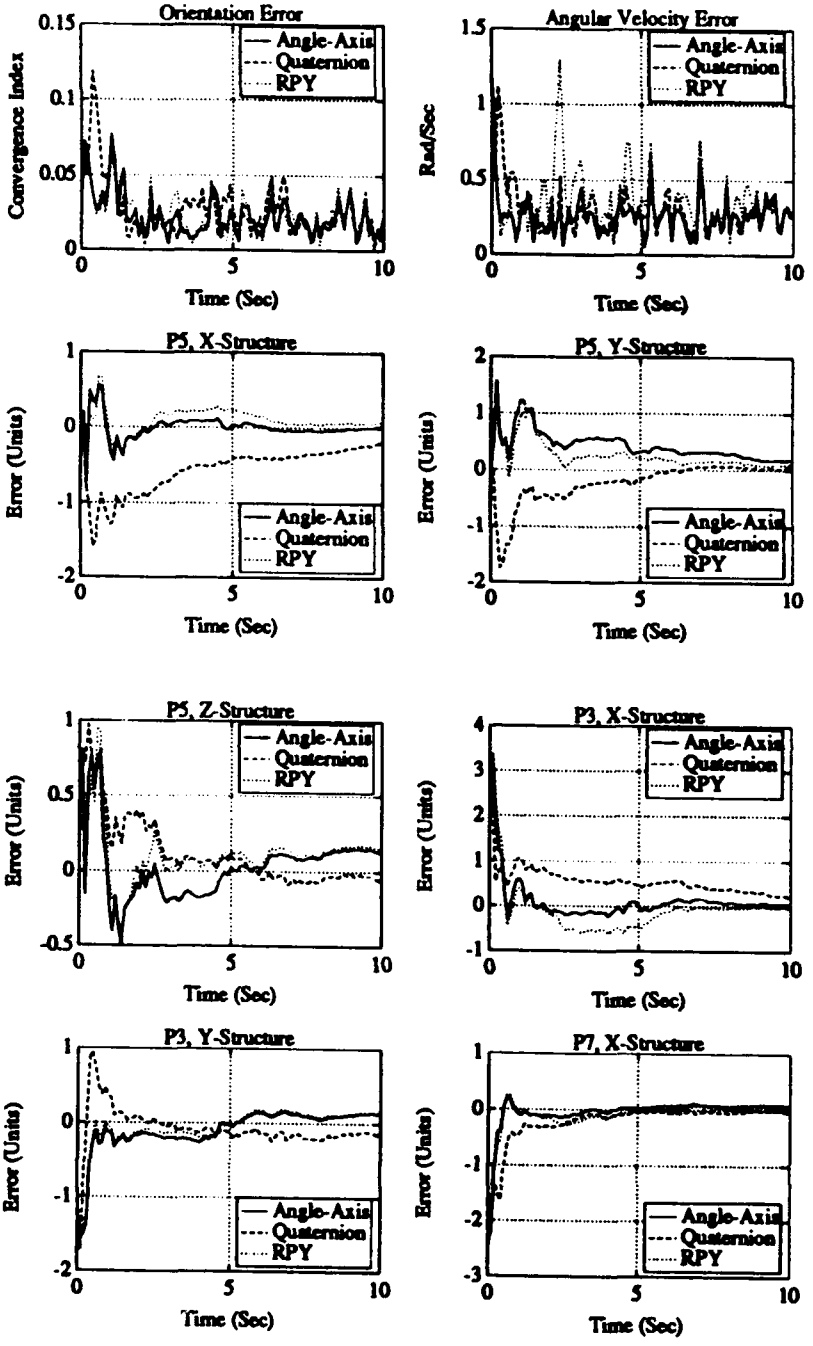

Figure 4.13

Sample results for Case 2, orientation, angular velocity and structure.

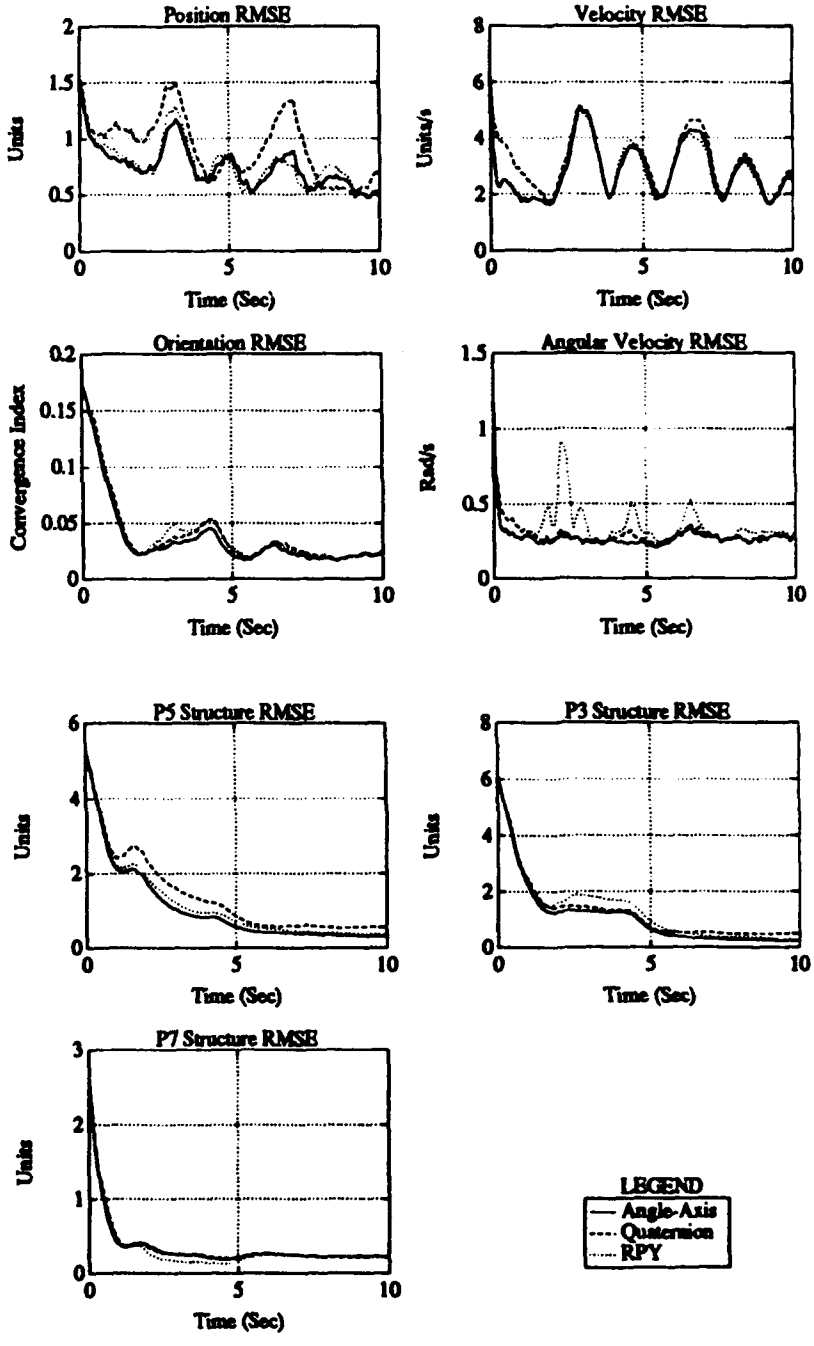

Figure 4.14

Monte Carlo simulation results for Case 2.

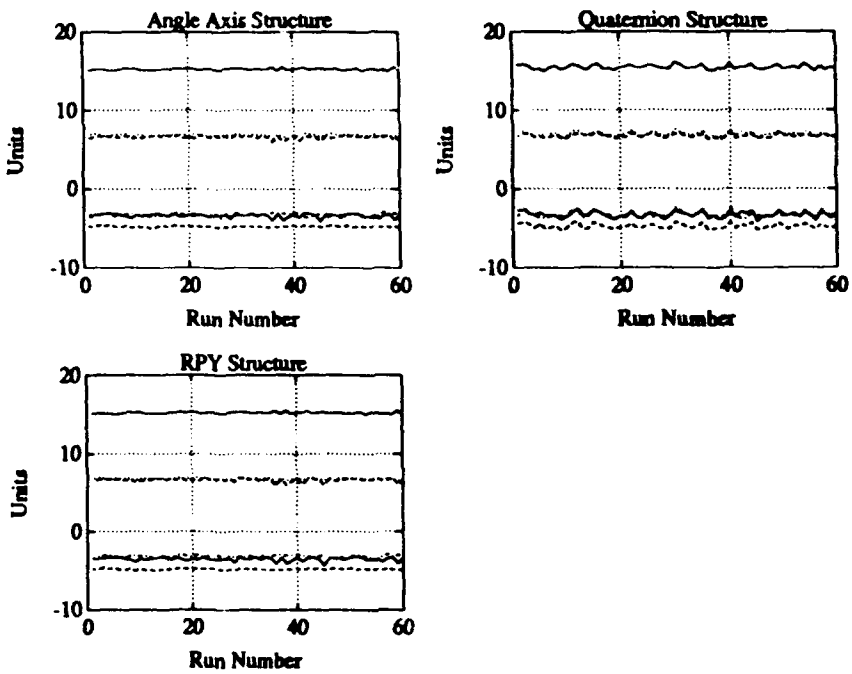

Figure 4.15

Final structure estimates in Monte Carlo simulations for Case 2.

Case 3: This trajectory defines true translational motion with constant acceleration of the block centroid, as in Case 2, but rotational motion is computed from a sinusoidal roll-pitch-yaw formulation. Again, let F_B denote the block-centred reference frame used in the imagery generation program with $\mathbf{I}_B^O \equiv \mathbf{I}_3$. The origin, O_B, of F_B lies at the block centroid, and the coordinate planes of F_B are parallel to faces of the block.

As in Case 2, translational motion of O_B was defined with

$$\begin{aligned} [\mathbf{T}_B^E]_E(0) &= [30, -20, 0]_E^T \text{ units}, \\ [\dot{\mathbf{T}}_B^E]_E(0) &= [2, 2, -2]_E^T \text{ units/s, and} \\ [\ddot{\mathbf{T}}_B^E]_E &= [-0.4, 0.4, 0.4]_E^T \text{ units/s}^2. \end{aligned} \tag{4.16}$$

However, translational motion of the object-centred frame is significantly different from that of Case 2 due to alternate rotational motion of F_B about O_B which is defined with sinusoidal variation of roll and pitch angles and linear yaw angle with time:

$$\phi(t) = a\sin(bt)$$

$$\theta(t) = a[1 - \cos(bt)] \tag{4.17}$$

$$\psi(t) = bt, \tag{4.18}$$

where

$$a = \frac{\pi}{6} \text{ and } b = \frac{2\pi}{10}. \tag{4.19}$$

Initialization parameters in this Case are identical to those for Case 2 shown in Table II. Figure 4.16 shows a subsequence of images from the left camera for this trajectory. The top of the object containing feature points 3 and 7 begins to rotate towards the observer (increasing pitch angle), while the right side containing points 1 and 3 appears to move down in the image due to positive yaw and roll angles. Again, the object returns to its original orientation at the end of the trajectory.

Sample results for initialization parameters given in Table II are shown in Figures 4.17 and 4.18. All three filters reliably track object position, but with noisier results for z-velocity than shown in Case 2. Almost identical velocity results are obtained from the three filters and, although noisy, do not show significant lag present in results for Case 2. The maximum gain strength, as in Case 2, is again maintained between 90 and 100 after initial transients. Very similar behavior between results for this trajectory and that of Case 2 is shown in Figure 4.18 for orientation and angular velocity estimation. Some bias, however, appears in structure estimates, particularly for the quaternion filter in p^5 x- and z-coordinates and p^3 y-coordinate.

Figure 4.19 shows Monte Carlo simulation results for this trajectory. As in Case 2, the quaternion filter gives slightly poorer performance in estimation of position and some structural components. Velocity root-mean-square errors lie below 2 units/s

during the second half of the trajectory and do not show significant oscillations that were present in results of Case 2. Orientation errors lie below 5% after about 2 seconds, with final errors of about 2.5% for all three filters, which is very similar to results of Case 2. Angular velocity errors decrease rapidly during the first 1 second and remain below about 3 rad/s for the remainder of the trajectory. Final root-mean-square errors for structure estimation are on the order of 0.5 to 1 unit.

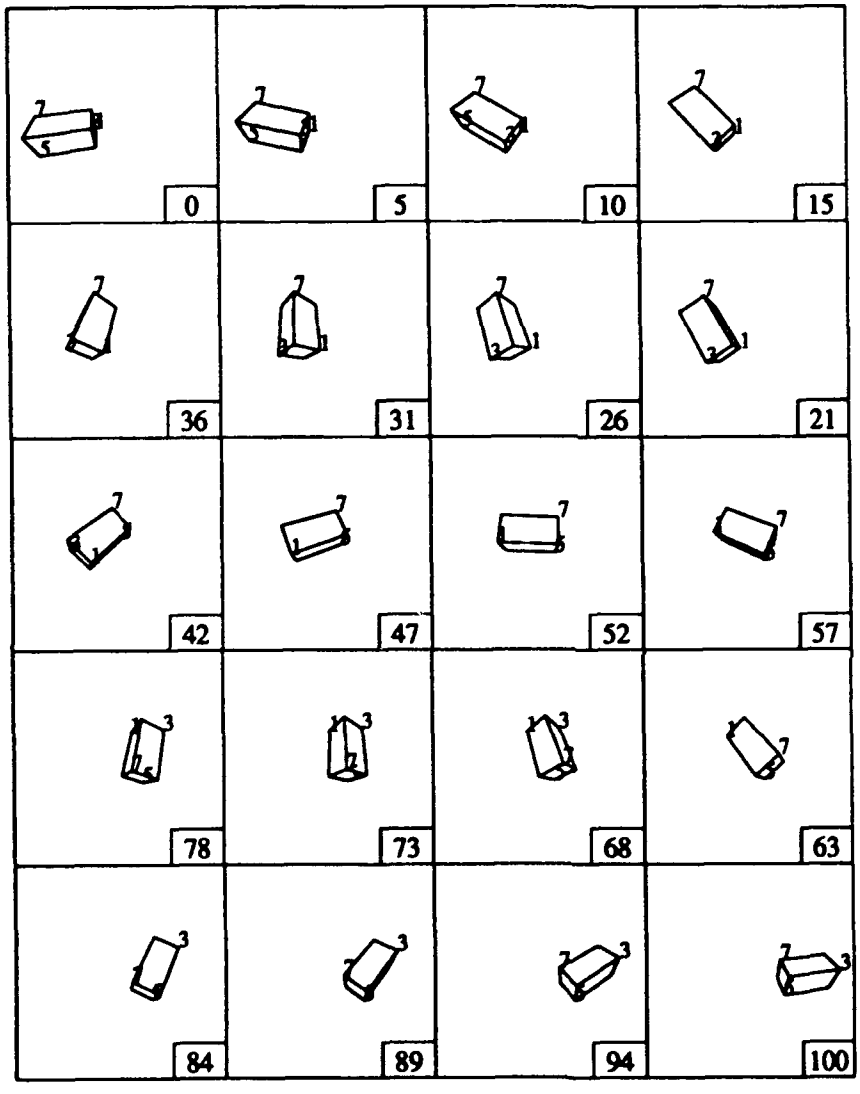

Figure 4.16

Image subsequence from the left camera for Case 3. Image frame numbers are shown in the lower left corner of each image. Four feature points, p^1, p^3, p^5, and p^7, are labelled in each image.

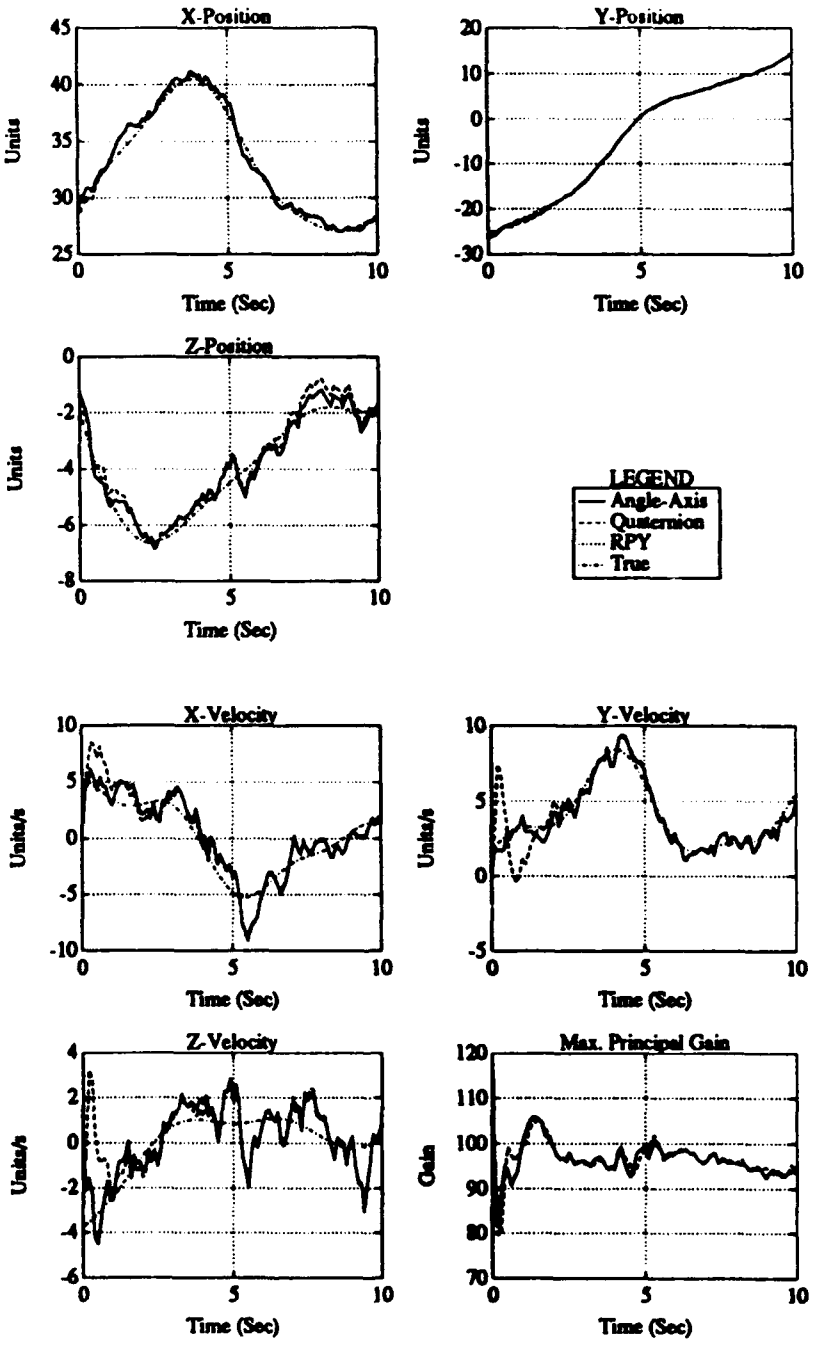

Figure 4.17

Sample results for Case 3, position, velocity and gain strength.

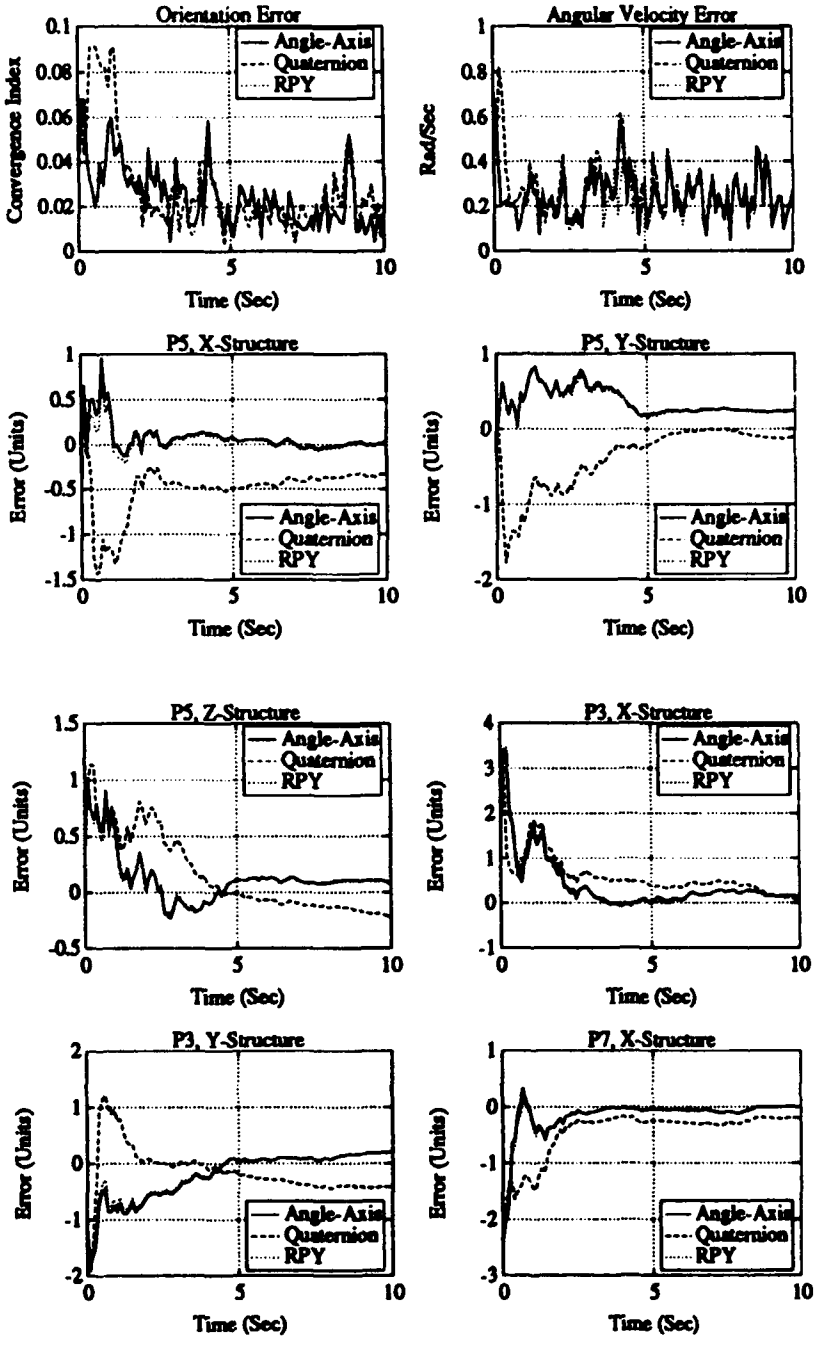

Figure 4.18

Sample results for Case 3, orientation, angular velocity and structure.

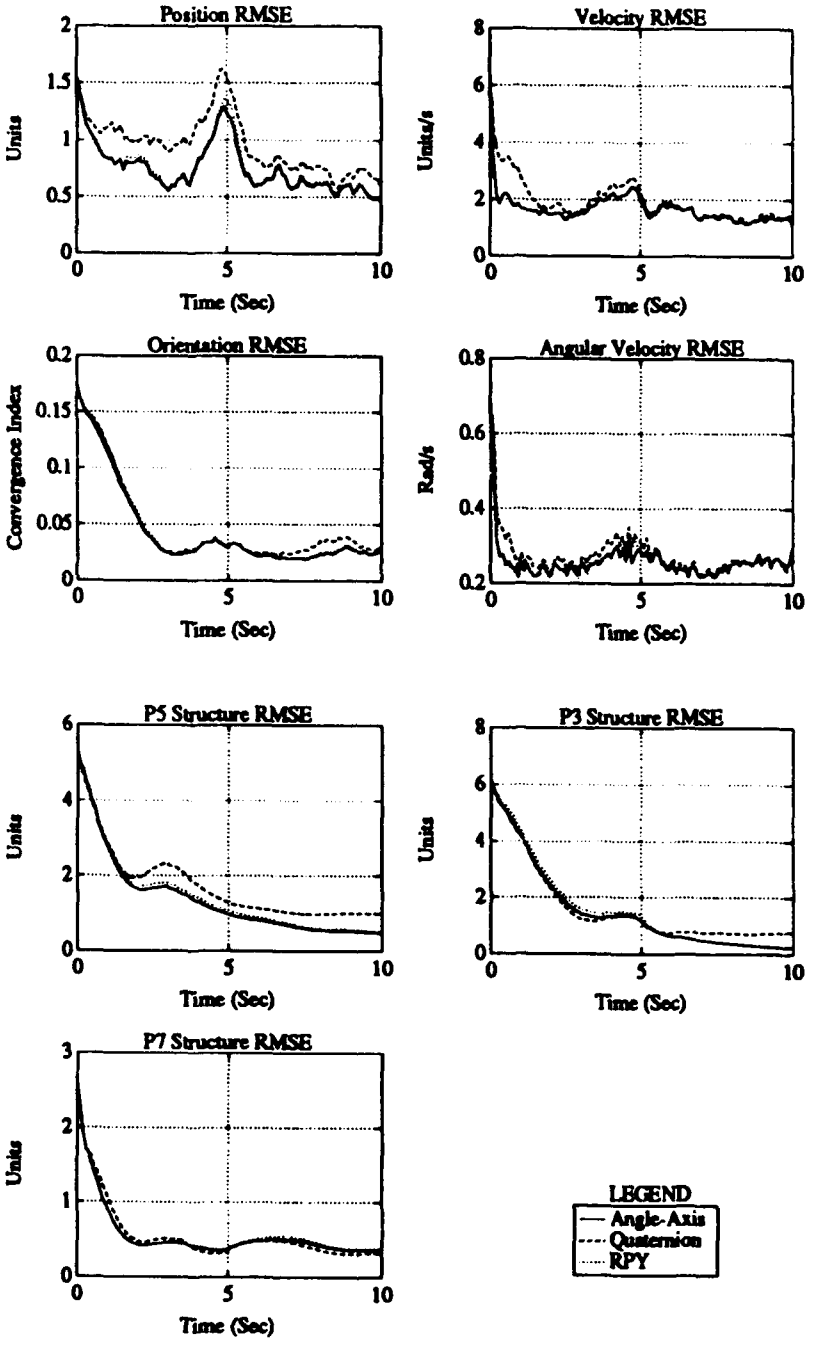

Figure 4.19

Monte Carlo simulation results for Case 3.

4.4 Chapter Summary

This Chapter has investigated approximate Cramer Rao bounds for a simple trajectory, discussed parameter selection for dynamic models, and presented a reduced set of simulation results which are representative of performance achieved for a wide range of trajectories. Cramer Rao bounds suggest very similar optimal performance for the three filters over a simple trajectory, and this is supported by sample and Monte Carlo simulation results for more complex trajectories presented in Section 4.3. One significant advantage of employing dynamic models developed in Chapter 3 is that model parameters can be selected as proposed by Singer [71] based to a large extent on the expected class of objects or motion, rather than purely through heuristic methods.

All three filters demonstrated good performance in tracking position and orientation of manoeuvring objects, however, estimation of translational velocity can exhibit filter lag, as in the trajectory of Case 2. Lag in the filters could be reduced to a certain extent by increasing the strength of filter gains (by increasing σ_t) but this may be in conflict with required measurement noise suppression. Angular velocity root-mean-squared errors for Cases 1 and 2 decrease very rapidly during the first 10 measurement events, and thereafter appear to remain at levels of about 0.3 rad/s ($\approx$ 17 Deg/s). In Case 2, angular velocity estimates from the roll-pitch-yaw filter showed distinct peaks in root-mean-squared errors which are due in part to some lag and overshoot in estimation of time derivatives of a sinusoidal pitch angle. Similar behavior, only on a smaller scale, is seen in Monte Carlo simulation results for Case 2 (see Figure 4.19) with a slight increase in angular velocity root-mean-squared errors just prior to 5 seconds. Generally, performance of the extended Kalman filters was degraded in the estimation of time derivatives of sinusoidally varying states.

Structure estimation for all three filters showed a sharp decrease in root-mean-squared errors during the first 1 or 2 seconds, with final accuracy after 100 observations of about 0.25 to 0.5 units. Structure estimates for the special feature point (p^7 for these simulations) for which the filters must estimate only a single component showed rapid convergence to within about 0.5 units after only 10 to 20 measurement events, while those for the feature point (p^3 here) in which two components must be estimated required about 50 measurement events.

5. Conclusions and Future Work

Many existing methods for motion and structure estimation proposed in the literature have been developed and examined for specific cases of translational and rotational motion. Violation of these assumptions often results in unobservability due to strict dependence of structural models on the nature of assumed motion of the object. The planar motion, constant velocity case of Chapter 2 falls into this category of strict motion assumptions, but provided a simplified framework to introduce the Kalman filtering approach while at the same time treating an important problem in motion analysis. The planar motion filter has demonstrated good performance over a wide range of trajectories with estimation accuracy in all states approaching Cramer Rao bounds.

Methods for more general motion proposed and evaluated in this report focus in part on removing restrictive assumptions concerning the nature of object motion by developing motion, structure, and measurement models for a manoeuvring object which is observed with a multiple-camera imaging system. Strict assumptions were not imposed on the mode of translational or rotational motion except for "smoothness" conditions in the sense that parameters can be differentiated with respect to time. Object manoeuvres, being "smooth" and time correlated, are modelled as first-order Gauss-Markov processes for both translational and rotational motion. The structural model has been defined on the basis of observed feature points only and is independent of object motion. In this case, three or more feature points observed by two or more cameras over multiple frames give sufficient information to estimate both motion and structure of the manoeuvring object.

Translational motion models were identical for all three filters. Good position tracking performance was demonstrated by the angle-axis and roll-pitch-yaw filters, with only slight degradation of position estimation shown by the quaternion filter. If the object is undergoing rotational motion, it is unlikely that the origin of the object-centred frame will coincide with the centre of rotation. As a result, translational motion in the presence of rotation will probably have strong sinusoidal components. Estimation of temporal derivatives of position in this case can exhibit lag and overshoot which led to the poor performance in velocity estimation shown in Case 2. This result demonstrates the conflicting requirements of maintaining large enough gains to track manoeuvring objects while at the same time keeping the gains small

enough to suppress measurement noise. Gain adjustment is easily accomplished, however, through selection of the manoeuvre correlation time constants and estimation of variances of the manoeuvre variables.

Unit quaternions have been examined by many authors in the parameterization of rotational motion. Imposing a unit norm constraint on the estimated quaternion, however, is not easily incorporated into observers such as the extended Kalman filter. Applying impulsive normalization of the estimated quaternion following each measurement update—a common proposed intervention in quaternion-based observers—was found to contribute significantly to filter instability. Instead, the quaternion filter is initialized with a unit quaternion, and propagates the quaternion estimate over time with constant norm. A unit quaternion and appropriately scaled structure vectors are then extracted as output variables from the state estimate. Note, however, that a continuum of valid solutions exist for quaternion filter; the squared norm of the quaternion estimate maintained within the filter is free to vary inversely with magnitudes of estimated structure vectors. In simulations, the quaternion filter showed slightly degraded performance in the estimation of position and some structure parameters in comparison to the roll-pitch-yaw and angle-axis filters.

The roll-pitch-yaw parameterization has previously been reported to be poorly behaved, and can lead to computationally demanding implementations. The approach taken here led, instead, to a very simple filter which may perform as well as the quaternion and angle-axis filters over some trajectories. However, some difficulties were encountered in the estimation of angular velocity for the trajectory of Case 2. This may have been due in part to sinusoidal temporal behavior of the true pitch angle since it was generally observed, as in the case of translational motion, that the extended Kalman filter performed poorly in the estimation of time derivatives of sinusoidally varying states.

Published results using the angle-axis parameterization in recursive motion and structure estimation are not available in the literature. Modelling rotational motion as a first-order Gauss-Markov process in the orientation vector and its time derivatives led not only to a simple implementation, but also performance which was as least as good as the other two parameterizations for all trajectories considered. Although this filter has very efficient time propagation due to the linear dynamic model, computation of the measurement model Jacobian is slightly more complex than the other two, and an occasional reset of the orientation vector and its time derivatives is required to maintain the rotation angle to within $\pm\pi$.

In all three filters, there are four possible orientations for the object-centred reference frame with respect to the three special feature points selected at initialization. When structure was initialized with all zeros, for example, convergence of structure estimates was observed in one of the four possible orientations of the object-centred frame, the particular one of the four being dependent on initial filter transients. In simulations with this object and the given measurement noise levels it was found that

the simple single-frame initialization provided sufficient initial accuracy for structure to converge in the object-centred frame selected at initialization, although this need not always be the case. Reliable convergence of structure estimates to true structural parameters was observed during all simulations, with the quaternion filter giving slightly noisier results that the angle-axis and roll-pitch-yaw filters. All structural parameters are modelled as constants and hence are not driven by process noise. As a result, the filter gains for structure states eventually become small as information is accumulated in the filter. Decoupling of structure states from the filters is possible and recommended after predefined confidence levels have been attained.

Several important factors, such as the expected nature of object motion, computational complexity, and required accuracy, must be considered when selecting a parameterization of rotational motion for any particular application. Prior information concerning motion constraints, for example the planar motion problem of Chapter 2, may clearly indicate appropriate and often simplified parameterizations and dynamic models. For more general motion, however, results of this investigation suggest that the angle-axis filter may provide a computationally efficient and sufficiently accurate means to recover both structure and motion of a manoeuvring object.

Research leading to this report has focused on only a single component of the overall motion analysis problem. Future work will investigate further components of this problem in the context of a hierarchical structure such as that proposed in Section 1.2. Generally, the proposed system is composed of three coupled multi-target tracking systems with increasing levels of complexity: multiple features are tracked in the image planes; multiple features are tracked in three-dimensions; and multiple objects, composed of multiple features, are tracked in three dimensions with six degrees-of-freedom. In such a system, occlusion, and temporal and spatial correspondence as well as rigid object segmentation represent challenging problems which have received little attention in the literature.

Another primary consideration in future work involves further evaluation of the extended Kalman filtering approach in comparison to other nonlinear observers. Adaptive extended Kalman filtering, Lyapunov methods, transformation to nonlinear observer canonical form, nonlinear map inversion, and sliding mode observers are representative of recently proposed methods which may lead to more robust observers for recovery of object motion and structure from multiple-camera image sequences.

References

[1] A. Rosenfeld, "Survey - Image analysis and computer vision: 1990," *Computer Vision, Graphics, and Image Processing*, Vol. 53, No. 3, May, 1991, pp 322-365.

[2] W.N. Martin, J.K. Aggarwal, "Survey: Dynamic scene analysis," *Computer Graphics and Image Processing*, Vol. 7, 1978, pp 356-374.

[3] H. Nagel, "Analysis techniques for image sequences," in *Proceedings of the Fourth International Conference on Pattern Recognition*, Kyoto, Japan, Nov. 7-10, 1978, pp 186-211.

[4] H. Nagel, "Image sequences - Ten (octal) years - From phenomenology towards a theoretical foundation," in *Proceedings of the Eight International Conference on Pattern Recognition*, Paris, France, Oct. 27-31, 1986, pp 1174-1185.

[5] J.K. Aggarwal, "Motion and time-varying imagery - An Overview," in *Proceedings, Workshop on Motion: Representation and Analysis*, Charleston, S.C., May 7-9, 1986, pp 1-6.

[6] J.K. Aggarwal, L.S.Davis, W.N. Martin, "Correspondence processes in dynamic scene analysis," *Proceedings of the IEEE*, Vol. 69, No. 5, May 1981, pp 562-572.

[7] J.K Aggarwal, A. Mitiche, "Structure and motion from images: fact and fiction," in *Procedings, Third Workshop on Computer Vision: Representation and Control*, Oct. 13-16, Bellaire, Michigan, 1985, pp 127-128.

[8] B.G. Schunck, "The image flow constraint equation," *Computer Vision, Graphics, and Image Processing*, Vol. 35, 1986, pp 20-46.

[9] B.G. Schunck, "Image flow segmentation and estimation by constraint line clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-11, No. 10, Oct. 1989, pp 1010-1027.

[10] K. Prazdny, "On the information in optical flows," *Computer Vision, Graphics, and Image Processing*, Vol. 22, 1983, pp 239-259.

[11] A.M. Waxman, S. Ullman, "Surface structure and three-dimensional motion from image flow kinematics," *International Journal of Robotics Research*, Vol. 4, No. 3, Fall 1985, pp 72-94.

[12] K. Kanatini, "Structure and motion from optical flow under perspective projection," *Computer Vision, Graphics, and Image Processing*, Vol. 38, 1987, pp 122-146.

[13] B.K.P. Horn, B.G. Schunck, "Determining optical flow," *Artificial Intelligence*, Vol. 17, 1981, pp 185-203.

[14] G. Adiv, "Determining three-dimensional motion and structure from optical flow generated by several moving objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-7, No. 4, July 1985, pp 384-401.

[15] J.W. Roach, J.K. Aggarwall, "Determining the movement of objects from a sequence of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 6, Nov. 1980, pp 554-562.

[16] R.Y. Tsai, T.S. Huang, "Estimating three-dimensional motion parameters of a rigid planar patch III: Finite point correspondences and the three view problem," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-32, No. 2, Apr. 1984, pp 213-220.

[17] R.Y. Tsai, T.S. Huang, "Uniqueness and estimation of three dimensional motion parameters of rigid objects with curved surfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 1, Jan. 1984, pp 13-26.

[18] J. Fang, T.S. Huang, "Some experiments on estimating the 3-D motion paramters of a rigid body from two consecutive image frames," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 5, Sept. 1984, pp 545-554.

[19] J. Weng, T.S. Huang, N. Ahuja, "3-D motion estimation, understanding and prediction from noisy image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-9, No. 3, May, 1987, pp 370-389.

[20] J. Weng, N. Ahuja, T.S. Huang, "Optimal motion and structure estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR'89*, San Diego, California, June 4-8, 1989, pp 144-152.

[21] J. Weng, P. Cohen, N. Rebibo, "Motion and structure estimation from stereo image sequences," *IEEE Transactions on Robotics and Automation*, Vol. 8, No. 3, June, 1992, pp 362-382.

[22] J. Philip, "Estimation of three-dimensional motion of rigid objects from noisy observations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-13, No. 1, Jan. 1991, pp 61-66.

[23] R.P. Futrelle, G.C. Speckert, "Extraction of motion data by interactive image processing," in *Proceedings, IEEE Conference on Pattern Recognition and Image Processing*, Chicago, Illinois, May 31-June 2, 1978, pp 405-408.

[24] D.B. Gennery, "Tracking known three-dimensional objects," in *Proceedings, National Conference on Artificial Intelligence*, AAAI-82, Pittsburgh, Pennsylvania, Aug. 18-20, 1982, pp 13-17.

[25] J. Hallam, "Resolving observer motion by object tracking," in *Proceedings of the Eighth International Joint Conference on Artificial Intelligence*, Karlsruhe, West Germany, Aug. 8-12, 1983, pp 792-797.

[26] G.R. Legters, T.Y. Young, "A mathematical model for computer image tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-4, No. 6, Nov. 1982, pp 583-594.

[27] J. Stuller, G. Krishnamurthy, "Kalman filter formulation of low-level television image motion estimation," *Computer Vision, Graphics, and Image Processing*, Vol. 21, 1983, pp 169-204.

[28] T. Broida, R. Chellappa, "Estimation of object motion parameters from noisy images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 1, Jan. 1986, pp 90-99.

[29] T. Broida, R. Chellappa, "Kinematics and structure of a rigid object from a sequence of noisy images," in *Proceedings, Workshop on Motion: Representation and Analysis*, Charleston, S.C., May 7-9, 1986, pp 95-100.

[30] T. Broida, R. Chellappa, "Kinematics of a rigid object from a sequence of noisy images: a batch approach," in *Proceedings, Computer Vision and Pattern Recognition*, Miami Beach, Fla., June 22-26, 1986, pp 176-182.

[31] T. Broida, R. Chellappa, "Performance bounds for estimating three dimensional motion parameters from a sequence of noisy images," *Journal of the Optical Society of America*, Vol. 6, No. 6, June 1989, pp 879-889.

[32] T. Broida, S. Chandrashekhar, R. Chellappa, "Recursive 3-D motion estimation from a monocular image sequence," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 26, No. 4, July 1990, pp 639-656.

[33] G.J. Young, R. Chellappa, "3-D motion estimation using a sequence of noisy stereo images," in *Proceedings, Computer Vision and Pattern Recognition*, Ann Arbor, Michigan, June 5-9, 1988, pp 710-716.

[34] G.J. Young, R. Chellappa, "3-D motion estimation using a sequence of noisy stereo images: models, estimation, and uniqueness results," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-12, No. 8, Aug. 1990, pp 735-759.

[35] G.S. Young, R. Chellappa, T.H. Wu, "Monocular motion estimation using a long sequence of noisy images," in *Proceedings, IEEE International Conference on Acoustics, Speech, and Signal Processing*, Toronto, Ontario, May 14-17, 1991, pp 2437-2440.

[36] T.J. Broida, R. Chellappa, "Estimating the kinematics and structure of a rigid object from a sequence of monocular images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 6, 1991, pp 497-513.

[37] T.J. Broida, "Estimating the kinematics and structure of a moving object from a sequence of images," Ph.D. dissertation, University of Southern California, Los Angeles, 1987.

[38] J.J. Wu, R.E. Rink, T.M. Caelli, V.G. Gourishankar, "Recovery of the 3-D location and motion of a rigid object through camera image (an extended Kalman filter approach)," *International Journal of Computer Vision*, Vol. 3, 1988, pp 373-394.

[39] K. Chang, H. Lee, C. Chung, "Adaptive tracking system for a manoeuvring target using images with correlated noises," *International Journal of Systems Science*, Vol. 20, No. 5, 1989, pp 839-857.

[40] N. Ayache, O. Faugeras, "Maintaining representations of the environment of a mobile robot," *IEEE Transactions on Robotics and Automation*, Vol. 5, No. 6, Dec. 1989, pp 804-819.

[41] N. Ayache, F. Lustman, "Trinocular stereo vision for robotics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 1, Jan. 1991, pp 73-85.

[42] G. Randall, S. Foret, N. Ayache, "Final steps towards real time trinocular stereo vision," in *Proceedings of the First European Conference on Computer Vision, ECCV-90*, Antibes, France, April 23-27, 1990, pp 601-603.

[43] L. Matthies, T. Kanade, "The cycle of uncertainty and constraint in robot perception," in *Proceedings of the Fourth International Symposium on Robotics Research*, 1987, pp 327-335.

[44] L. Matties, T. Kanade, R. Szeliski, "Kalman filter-based algorithms for estimating depth from image sequences," *International Journal of Computer Vision*, Vol. 3, 1989, pp 209-236.

[45] E.D. Dickmanns, V. Graefe, "Dynamic monocular machine vision," *International Journal of Machine Vision and Applications*, Springer-Verlag, New York, Vol. 1, No. 4, 1988, pp 223-240.

[46] E.D. Dickmanns, V. Graefe, "Applications of dynamic monocular machine vision," *International Journal of Machine Vision and Applications*, Springer-Verlag, New York, Vol. 1, No. 4, 1988, pp 241-261.

[47] V. Graefe, "A pre-processor for the real-time interpretation of dynamic scenes," in T.S. Huang (Ed.), *Image Sequence Processing and Dynamic Scene Analysis*, Springer-Verlag, Berlin, 1983, pp 519-531.

[48] H.G. Meissner, E.D. Dickmanns, "Control of an unstable plant by computer vision," in T.S. Huang (Ed.), *Image Sequence Processing and Dynamic Scene Analysis*, Springer-Verlag, Berlin, 1983, pp 532-548.

[49] R. Schell, E.D. Dickmanns, "Autonomous automatic landing through computer vision," in *AGARD Conference Proceedings No. 455*, Guidance and Control Panel 48th Symposium, Instituto da Defesa Nacional in Lisbon, Portugal, May 9-12, 1989, pp 24:1-24:9.

[50] E.D. Dickmanns, B. Mysliwitz, T. Christians, "An integrated spatio-temporal approach to automatic guidance of autonomous vehicles," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 20, No. 6, Nov./Dec. 1990, pp 1273-1284.

[51] E.D. Dickmanns, T. Christians, "Relative 3D-state estimation for autonomous visual guidance of road vehicles," *Robotics and Autonomous Systems*, Vol. 7, 1991, pp 113-123.

[52] E.D. Dickmanns, B.D. Mysliwetz, "Recursive 3-D road and relative ego-state recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, Feb. 1992, pp 199-213.

[53] S.-L. IU, K. Wohn, "Estimation of general rigid body motion from a sequence of images," in *Proceedings of the IEEE 10th International Conference on Pattern Recognition*, Vol. 1, Atlantic City, NJ, USA, June 16-21, 1990, pp 217-219.

[54] S.-L. IU, K. Wohn, "Recovery of 3D motion of a single particle," *Pattern Recognition*, Vol. 24, No. 3, 1991, pp 241-252.

[55] B.K.P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society of America, A*, Vol. 4, No. 4, April 1987, pp 629-642.

[56] I.K. Sethi, R. Jain, "Finding trajectories of feature points in a monocular image sequence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, No. 1, Jan. 1987, pp 56-73.

[57] V. Salari, I.K. Sethi, "Feature point correspondence in the presence of occlusion," *IEEE Transactions of Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, Jan. 1990, pp 87-91.

[58] V.C. Aitken, "Motion and structure estimation in noisy monocular image sequences: specialization to planar motion," Master's Thesis, Carleton University, Ottawa, Ontario, 1991.

[59] V.C. Aitken, H.M. Schwartz, "Motion and structure from planar motion in monocular image sequences," in *Proceedings of the Third Conference on Military Robotic Applications, MRV91*, Medicine Hat, Alberta, Sept. 9-12, 1991, pp 177-184.

[60] R.E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions ASME, Ser. D: Journal of Basic Engineering*, Vol. 82, 1960, pp 35-45.

[61] J.H. Taylor, "The Cramer-Rao estimation error lower bound computation for deterministic nonlinear systems," *IEEE Transactions on Automatic Control*, Vol. AC-24, No.2, April 1979, pp 343-344.

[62] C. Chang, J. Tabaczynski, "Application of state estimation to target tracking," *IEEE Transactions on Automatic Control*, Vol. AC-29, No. 2, Feb. 1984, pp 98-109.

[63] M.H. Ang, V.D. Tourassis, "Singularities of Euler and roll-pitch-yaw representations," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-23, No. 3, May 1987, pp 317-324.

[64] I.Y. Bar-Itzhack, "Optimum normalization of a computed quaternion of rotation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-7, Mar. 1971, pp 401-402.

[65] I.Y. Bar-Itzhack, Y. Oshman, "Attitude determination from vector observations: quaternion estimation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-21, No. 1, Jan. 1985, pp 128-135.

[66] M. Tahk, J.L Spayer, "Target tracking problems subject to kinematic constraints," *IEEE Transactions on Automatic Control*, Vol. 35, No. 3, March, 1990, pp 324-326.

[67] J. Stuelpnagel, "On the parameterization of the three-dimensional rotation group," *SIAM Review*, Vol. 6, No. 4, Oct. 1964, pp 422-430.

[68] Y.L. Gu, "An exploration of orientation representation by Lie algebra for robotic applications," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 20, No. 1, Jan. 1990, pp 243-248.

[69] J.C.K. Chou, "Quaternion kinematic and dynamic differential equations," *IEEE Transactions on Robotics and Automation*, Vol. 8, No. 1, Feb. 1992, pp 53-64.

[70] A. Edwards, "The state of strapdown inertial quidance and navigation," *Journal of The Institute of Navigation*, Vol. 18, No. 4, Winter, 1971-72, pp 386-401.

[71] R.A. Singer, "Estimating optimal tracking filter performance for manned manoeuvring targets," *IEEE Transactions on Aerospace and Electrinic Systems*, Vol. AES-6, No. 4, July, 1970, pp 473-483.

[72] R. Hermann, A.J. Krener, "Nonlinear controllability and observability," *IEEE Transactions on Automatic Control*, Vol. AC-22, No. 5, Oct. 1977, pp728-740.

[73] S. Ullman, *The Interpretation of Visual Motion*, Cambridge, MA: M.I.T. Press, 1979.

[74] A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, New York: Academic, 1970.

[75] A. Gelb, *Applied Optimal Estimation*, Cambridge MA: M.I.T. Press, 1974.

[76] B.D.O. Anderson, J.B. Moore, *Optimal Filtering*, Englewood Cliffs, NJ: Prentice-Hall, 1979.

[77] P.S. Maybeck, *Stochastic Models, Estimation, and Control*, New York: Academic, Vol. 1, 1979, and Vol. 2, 1982.

[78] A.V. Balakrishnan, *Kalman Filtering Theory*, Optimization Software Inc., Publications Division, New York, 1984.

[79] C.R. Rao, *Linear Statistical Inference and Its Applications, Second Edition*, New York: John Wiley and Sons, 1973.

[80] K.V. Bury, *Statistical Models in Applied Science*, New York: John Wiley and Sons, 1975.

[81] N.U. Ahmed, *Elements of Finite-Dimensional Systems and Control Theory*, Pittman Monographs and Surveys in Pure and Applied Mathematics, 37, Longman Scientific and Technical: Harlow Essex, England, 1988.

[82] B. Etkin, *Dynamics of Atmospheric Flight*, New York: John Wiley and Sons, 1972.

[83] V.N. Branets and I.P. Shmyglevski, *Primeneniye Kvaternionov V Zadachakh Orientatsii Tverdogo Tela*, (Application of Quaternions To Rigid Body Rotation Problems), Nauka Press, Moscow, 1973, NASA technical translation NASA TT F-15,414, Translated for National Aeronautics and Space Administration under contract No. NASw 2483, by SCITRAN, P.O. Box 5456, Santa Barbera, California, 93108.

[84] P.C. Hughes, *Spacecraft Attitude Dynamics*, New York: John Wiley and Sons, 1986.

Appendix A
Parameterizations of Relative Orientation

Chapters 2 and 3 introduced notation and defined a change-of-basis transformation $\mathbf{I}_E^O$ such that a vector $\mathbf{r}_O$ in F_O is transformed to $[\mathbf{r}_O]_E$, its representation in F_E, through the expression

$$[\mathbf{r}_O]_E = \mathbf{I}_E^O \mathbf{r}_O. \tag{A.1}$$

As before, the object-centred frame, F_O, is defined by a point O_O, the origin of F_O, and a right-handed, ordered, orthonormal basis set $\{\mathbf{i}_O, \mathbf{j}_O, \mathbf{k}_O\}$. Equivalent notation also applies to the earth-fixed frame F_E. In order to focus on relative orientation only, it is assumed throughout this appendix that O_O and O_E are coincedent. In this case,

$$[\mathbf{r}_O]_E \equiv \mathbf{r}_E \tag{A.2}$$

since the translation vector from F_E to F_O is zero. The orthogonal transformation $\mathbf{I}_E^O$ belongs to the three-dimensional rotation group, denoted here by $\mathcal{R}_3$, for which there are about eight commonly used representations [67]. This appendix provides a tutorial on three such representations: the Euler angle-axis, quaternion (sometimes called Euler parameters, or Rodriquez-Hamilton parameters), and roll-pitch-yaw (sometimes called Euler angles) parameterizations. The material presented in this appendix draws primarily from references [67]-[70], [82]-[84].

The problem of parameterizing the group $\mathcal{R}_3$ has been of interest since 1776, when Euler first showed that this group is itself a three-dimensional manifold. Euler observed that the general displacement of a rigid body with one fixed point is a rotation about an axis through that point (Euler's theorem). In the present context, Euler's theorem states that F_O is related to F_E by a rigid rotation of $\mathbb{R}^3$ about an axis through their common origin. Since $\mathcal{R}_3$ is known to be three-dimensional, the rotational degrees-of-freedom in $\mathbb{R}^3$ are at most three. However, it is topologically impossible to have a global and nonsingular three-dimensional parameterization of $\mathcal{R}_3$, [67]. In fact, Hopf showed in 1940 that the minimum number of parameters required to represent $\mathcal{R}_3$ in a one-to-one global manner is five, (see [67]). The quaternion representation, a four-dimensional parameterization of $\mathcal{R}_3$, results in a two-to-one mapping of $\mathbb{R}^4$ onto $\mathcal{R}_3$, and is often considered sufficient for most practical purposes.

It is important to appreciate a rather subtle distinction between two interpretations of operations performed by elements of $\mathcal{R}_3$. One interpretation treats an element $\mathbf{R} \in \mathcal{R}_3$ as a *rotation* operator which relates two vectors, say $\mathbf{x}$ and $\mathbf{y}$ with $\|\mathbf{x}\| = \|\mathbf{y}\|$, in the same bases according to $\mathbf{y} = \mathbf{R}\mathbf{x}$. In this case, $\mathbf{R}$ rotates the vector $\mathbf{x}$ into a new vector $\mathbf{y}$. The second interpretation treats an element $\mathbf{I}_\alpha^\beta \in \mathcal{R}_3$ as an orthogonal (or identity, hence the use of 'I') *change-of-basis transformation* which operates on a vector $\mathbf{r}_\beta$ expressed with respect to F_β to obtain a representation for the same vector, $\mathbf{r}_\alpha$, expressed with respect to F_α according to $\mathbf{r}_\alpha = \mathbf{I}_\alpha^\beta \mathbf{r}_\beta$. The main point is that vectors $\mathbf{r}_\beta$ and $\mathbf{r}_\alpha$ represent the same physical entity, while the vectors $\mathbf{x}$ and $\mathbf{y}$ in general do not. It is often the case that "rotation" is used interchangeably to imply either interpretation. This appendix deals exclusively with the "change-of-bases transformation" interpretation.

This review first introduces fundamental properties of rotation matrices in Section A.1, treats the Euler angle-axis parameterization in Section A.2, quaternions in Section A.3, and the roll-pitch-yaw representation in Section A.4.

A.1 Fundamental Properties of $\mathbf{I}_E^O$

The following list states some fundamental and well-known properties of the matrix $\mathbf{I}_E^O$ representing the change-of-bases transformation from F_O to F_E.

Property 1 The matrix $\mathbf{I}_E^O$ is orthogonal with determinant $+1$.

Property 2 The matrix $\mathbf{I}_E^O$ has eigenvalues $\{+1, e^{\pm j\gamma}\}$. Euler's axis of rotation of $\mathbf{I}_E^O$ is parallel to the (normalized) eigenvector, $\bar{\xi}$, corresponding to the eigenvalue $+1$, while the angle of rotation, γ, about $\bar{\xi}$ is measured with positive sense from F_E to F_O according to the right-hand-rule about $\bar{\xi}$.

Property 3 The matrix $\mathbf{I}_E^O$ can be written as

$$\begin{aligned} \mathbf{I}_E^O &= \left[\, [\mathbf{i}_O]_E \;\; [\mathbf{j}_O]_E \;\; [\mathbf{k}_O]_E \, \right] \\ &= \left[\, [\mathbf{i}_E]_O \;\; [\mathbf{j}_E]_O \;\; [\mathbf{k}_E]_O \, \right]^T, \end{aligned} \tag{A.3}$$

where superscript 'T' denotes transposition.

Property 4 The angular velocity of F_O with respect to F_E expressed in F_E, denoted $[\mathbf{w}_E^O]_E$, can be computed by noting that

$$\begin{aligned} \frac{\mathrm{d}[\mathbf{i}_O]_E}{\mathrm{d}t} &= [\mathbf{w}_E^O]_E \times [\mathbf{i}_O]_E, \\ \frac{\mathrm{d}[\mathbf{j}_O]_E}{\mathrm{d}t} &= [\mathbf{w}_E^O]_E \times [\mathbf{j}_O]_E, \\ \frac{\mathrm{d}[\mathbf{k}_O]_E}{\mathrm{d}t} &= [\mathbf{w}_E^O]_E \times [\mathbf{k}_O]_E, \end{aligned} \tag{A.4}$$

which gives

$$\dot{\mathbf{I}}_E^O = [\mathbf{w}_E^O]_E^{\times} \mathbf{I}_E^O, \tag{A.5}$$

where $[\cdot]^{\times}$ is the matrix cross-product operator defined in terms of components of a vector $\xi = [\bar{\xi}_1, \bar{\xi}_2, \bar{\xi}_3]^T$ as

$$\xi^{\times} \triangleq \begin{bmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{bmatrix}. \tag{A.6}$$

Rearranging (A.5) gives

$$[\mathbf{w}_E^O]_E^{\times} = \dot{\mathbf{I}}_E^O \left(\mathbf{I}_E^O\right)^T = \dot{\mathbf{I}}_E^O \mathbf{I}_O^E. \tag{A.7}$$

The matrix (tensor) $[\mathbf{w}_E^O]_E^{\times}$ is basis-dependent and is transformed to F_O according to

$$\begin{aligned} [\mathbf{w}_E^O]_O^{\times} &= \mathbf{I}_O^E [\mathbf{w}_E^O]_E^{\times} \mathbf{I}_E^O \\ &= \mathbf{I}_O^E \dot{\mathbf{I}}_E^O \mathbf{I}_O^E \mathbf{I}_E^O \\ &= \mathbf{I}_O^E \dot{\mathbf{I}}_E^O \end{aligned} \tag{A.8}$$

Property 1 follows from the fact that both F_O and F_E are right-handed orthonormal triads. Property 2 then follows immediately from Property 1. Property 3 is derived by considering identities of the form $[\mathbf{i}_O]_E = \mathbf{I}_E^O [\mathbf{i}_O]_O$ and noting that $[\mathbf{i}_O]_O = [1, 0, 0]^T$. Property 4 is fundamental in derivations of relationships between angular velocity and the selected parameter set and their time derivatives.

A.2 Euler Angle-Axis Parameterization

According to Euler's theorem, a rigid rotation of $\mathbb{R}^3$ is completely defined by a unit vector $\bar{\xi}$ which specifies the axis of the basis transformation operator $\mathbf{I}_E^O$, and an angle γ through which the coordinate system is rotated about $\bar{\xi}$. This is a four-parameter representation with one quadratic constraint. Note that $\bar{\xi}$ is invariant under the basis transformation and hence the subscript specifying the frame to which $\bar{\xi}$ is referenced has been dropped.

Given an angle γ and axis $\bar{\xi}$ the corresponding matrix $\mathbf{I}_E^O$ can be determined by the following argument. Suppose at $t = 0$ reference frames F_O and F_E are aligned so that $\mathbf{I}_E^O(t = 0) = \mathbf{I}_3$, the 3×3 identity matrix. At time $t = 1$, the orientation of reference frame F_O relative to F_E is given by a rotation about $\bar{\xi}$ in F_E through an angle γ. For the purposes of deriving $\mathbf{I}_E^O(t = 1)$, the motion during this time interval can be modelled with constant angular velocity $[\mathbf{w}_E^O]_E = \gamma\bar{\xi}$. Integration of (A.5)

from $t = 0$ to $t = 1$ with initial condition $\mathbf{I}_E^O(t = 0) = \mathbf{I}_3$ immediately gives the matrix exponential representation

$$\mathbf{I}_E^O(t = 1) = \exp\left(\gamma\bar{\xi}^{\mathbb{I}}\right), \tag{A.9}$$

or, with

$$\xi \triangleq \gamma\bar{\xi}, \tag{A.10}$$

the basis transformation in terms of ξ is

$$\mathbf{I}_E^O(\xi) = \exp\left(\xi^{\mathbb{I}}\right). \tag{A.11}$$

Equation (A.11) is a three-dimensional parameterization in which ξ is often called the *orientation vector*. This representation has been used extensively by Faugeras [40], for example. The matrix cross product operator, $(\cdot)^{\mathbb{I}}$, yields a 3×3 skew symmetric matrix. As suggested by (A.9), every 3×3 rotation matrix can be expressed as an exponential function of some 3×3 skew symmetric matrix, which has led to elegant studies in terms of Lie algebra as an approach to resolving the problem of rotation and orientation representations, [68].

By applying the well-known Cayley-Hamilton theorem to the characteristic equation of the skew symmetric matrix $\bar{\xi}^{\mathbb{I}}$, and using the fact that $\bar{\xi}$ is a unit vector, the following important identity is obtained:

$$\left(\bar{\xi}^{\mathbb{I}}\right)^3 + \bar{\xi}^{\mathbb{I}} = 0. \tag{A.12}$$

A result of (A.12) is that the highest order of any power series of $\bar{\xi}^{\mathbb{I}}$ is two. Consequently, (A.9) can be written as

$$\begin{aligned} \mathbf{I}_E^O(\gamma, \bar{\xi}) &= \sum_{j=0}^{\infty} \frac{1}{j!}\left(\gamma\bar{\xi}^{\mathbb{I}}\right)^j \\ &= \mathbf{I}_3 + \sin(\gamma)\bar{\xi}^{\mathbb{I}} + [1 - \cos(\gamma)]\left(\bar{\xi}^{\mathbb{I}}\right)^2. \end{aligned} \tag{A.13}$$

With the identity

$$\left(\bar{\xi}^{\mathbb{I}}\right)^2 \equiv \bar{\xi}\bar{\xi}^T - \mathbf{I}_3, \tag{A.14}$$

the result in Equation (A.13) can also be written as

$$\mathbf{I}_E^O(\gamma, \bar{\xi}) = \cos(\gamma)\mathbf{I}_3 + \sin(\gamma)\bar{\xi}^{\mathbb{I}} + [1 - \cos(\gamma)]\bar{\xi}\bar{\xi}^T. \tag{A.15}$$

The angle of rotation, γ, can be obtained from $\mathbf{I}_E^O$ either by using Eigen analysis with Property 2 of Section A.1 and recalling that the trace of a matrix equals the sum of its eigenvalues, or by noting that in (A.13) $\bar{\xi}^{\mathbb{I}}$ is skew symmetric and[1] $\mathrm{Tr}\{(\bar{\xi}^{\mathbb{I}})^2\} = -2$. In either case,

$$\cos(\gamma) = \frac{1}{2}\left(\mathrm{Tr}\{\mathbf{I}_E^O\} - 1\right). \tag{A.16}$$

[1]The notation $\mathrm{Tr}\{\mathbf{A}\}$ denotes the trace of the matrix $\mathbf{A}$.

In addition, the axis $\bar{\xi}$ may be obtained from (A.13) with

$$\bar{\xi}^{\mathbb{I}} = \frac{1}{2\sin(\gamma)}\left(\mathbf{I}_E^O - (\mathbf{I}_E^O)^T\right). \tag{A.17}$$

Because a rotation about $\bar{\xi}$ through an angle γ is equivalent to a rotation about $-\bar{\xi}$ through an angle $-\gamma$, the angle of rotation may be restricted to the interval $0 \leq \gamma \leq \pi$ for unique inversion of the cosine in (A.16). In this case, $\bar{\xi}$ computed from (A.17) together with γ define the transformation $\mathbf{I}_E^O$. However, numerical difficulties may arise when $\mathbf{I}_E^O$ is symmetric ($\gamma = 0,\ \pi$) for which the expression in (A.17) is of indeterminate form (0/0) indicating the mathematical singularities of this parameterization; when $\gamma = 0$ the axis $\bar{\xi}$ is undefined and when $\gamma = \pi$ the sign of $\bar{\xi}$ is arbitrary.

It is important to note the difference between the axis, $\bar{\xi}$, of the basis transformation operator, $\mathbf{I}_E^O$, and what is commonly called the "axis of rotation" or instantaneous angular velocity of the object $\mathbf{w}(t) \triangleq [\mathbf{w}_E^O]_E(t)$. In the following development, all time derivatives are with respect to an observer in the earth-fixed frame, and, although the subscript 'E' is omitted, $\dot{\bar{\xi}}$ is expressed in F_E. Several identities are required in the derivation of $\mathbf{w}$, and are also used to write $\dot{\gamma}$ and $\dot{\bar{\xi}}$ in terms of $\mathbf{w}$:

1. Because $\bar{\xi}$ is by definition a unit vector,

$$\bar{\xi}^T\bar{\xi} = 1, \text{ and} \tag{A.18}$$

$$\dot{\bar{\xi}}^T\bar{\xi} = \bar{\xi}^T\dot{\bar{\xi}} = 0, \tag{A.19}$$

   which means that $\bar{\xi}$ is always orthogonal to its time derivative.

2. For any vector $\mathbf{a}$, $\dot{\bar{\xi}} \times (\bar{\xi} \times \mathbf{a})$ is collinear with $\bar{\xi}$ and hence

$$\bar{\xi} \times (\dot{\bar{\xi}} \times (\bar{\xi} \times \mathbf{a})) = 0. \tag{A.20}$$

   In terms of matrix cross product operators, (A.20) is written as

$$\bar{\xi}^{\mathbb{I}}(\dot{\bar{\xi}})^{\mathbb{I}}\bar{\xi}^{\mathbb{I}} = \mathbf{0}. \tag{A.21}$$

3. The set $\{\bar{\xi}, \dot{\bar{\xi}}, \bar{\xi}^{\mathbb{I}}\dot{\bar{\xi}}\}$ forms a right-handed orthogonal triad and hence

$$\bar{\xi}^T\bar{\xi}^{\mathbb{I}}\dot{\bar{\xi}} = 0, \text{ and} \tag{A.22}$$

$$\bar{\xi}^{\mathbb{I}}\bar{\xi}^{\mathbb{I}}\dot{\bar{\xi}} = -\dot{\bar{\xi}}. \tag{A.23}$$

   Similar identities can be applied to various combinations of the three elements of this triad.

4. Two standard results from vector analysis give, for any $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$,

$$\begin{aligned} \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) &= \mathbf{a}^T\mathbf{c}\mathbf{b} - \mathbf{a}^T\mathbf{b}\mathbf{c}, \text{ and} \\ (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} &= (\mathbf{b}\mathbf{a}^T - \mathbf{a}\mathbf{b}^T)\mathbf{c}. \end{aligned} \tag{A.24}$$

Substitution of (A.13) into (A.7) followed by repeated applications of identities (A.12) and (A.21) results in

$$\mathbf{w}^{\shortparallel} = \dot{\gamma}\bar{\xi}^{\shortparallel} + \sin\gamma\dot{\bar{\xi}}^{\shortparallel} + (1-\cos\gamma)\left(\bar{\xi}^{\shortparallel}\dot{\bar{\xi}}^{\shortparallel} - \dot{\bar{\xi}}^{\shortparallel}\bar{\xi}^{\shortparallel}\right). \tag{A.25}$$

All terms in (A.25) are cross product operators. With identities (A.24) and using the obvious result that $\bar{\xi}^{\shortparallel}\bar{\xi} = \mathbf{0}$, the following result holds for any vector $\mathbf{c} \in \mathbb{R}^3$:

$$\begin{aligned}\left(\bar{\xi}^{\shortparallel}\dot{\bar{\xi}}^{\shortparallel} - \dot{\bar{\xi}}^{\shortparallel}\bar{\xi}^{\shortparallel}\right)\mathbf{c} &= \bar{\xi}\times(\dot{\bar{\xi}}\times\mathbf{c}) - \dot{\bar{\xi}}\times(\bar{\xi}\times\mathbf{c}) \\ &= (\dot{\bar{\xi}}\bar{\xi}^T - \bar{\xi}\dot{\bar{\xi}}^T)\mathbf{c} \\ &= (\bar{\xi}\times\dot{\bar{\xi}})\times\mathbf{c},\end{aligned} \tag{A.26}$$

which implies that

$$\left(\bar{\xi}^{\shortparallel}\dot{\bar{\xi}}^{\shortparallel} - \dot{\bar{\xi}}^{\shortparallel}\bar{\xi}^{\shortparallel}\right) = \left(\bar{\xi}^{\shortparallel}\dot{\bar{\xi}}\right)^{\shortparallel}. \tag{A.27}$$

This result allows the cross product operators of (A.25) to be removed which gives,

$$\mathbf{w} = \dot{\gamma}\bar{\xi} + \sin\gamma\dot{\bar{\xi}} + (1-\cos\gamma)\bar{\xi}^{\shortparallel}\dot{\bar{\xi}}. \tag{A.28}$$

Equation (A.28) demonstrates that the angular velocity has components on three orthogonal axis $\bar{\xi}$, $\dot{\bar{\xi}}$, and $\bar{\xi}^{\shortparallel}\dot{\bar{\xi}}$.

Multiplying both sides of (A.28) on the left by $\bar{\xi}^T$ and using identities (A.22) and its counterparts immediately gives

$$\dot{\gamma} = \bar{\xi}^T\mathbf{w}. \tag{A.29}$$

Multiplying both sides of (A.28) on the left by $\bar{\xi}^{\shortparallel}$ and using identity (A.23) gives

$$\bar{\xi}^{\shortparallel}\mathbf{w} = -(1-\cos(\gamma))\dot{\bar{\xi}} + \sin(\gamma)\bar{\xi}^{\shortparallel}\dot{\bar{\xi}}. \tag{A.30}$$

Multiplying both sides of (A.30) on the left by $\bar{\xi}^{\shortparallel}$ and again using identity (A.23) gives

$$\bar{\xi}^{\shortparallel}\bar{\xi}^{\shortparallel}\mathbf{w} = -\sin(\gamma)\dot{\bar{\xi}} - (1-\cos(\gamma))\bar{\xi}^{\shortparallel}\dot{\bar{\xi}}. \tag{A.31}$$

Equations (A.30) and (A.31) can be treated as two equations in the two unknowns $\dot{\bar{\xi}}$ and $\bar{\xi}^{\shortparallel}\dot{\bar{\xi}}$. Solving (A.30) for the quantity $\bar{\xi}^{\shortparallel}\dot{\bar{\xi}}$ and substituting the result into (A.30), followed by algebraic manipulation and the use of standard trigonometric identities gives

$$\dot{\bar{\xi}} = -\frac{1}{2}\left\{\bar{\xi}^{\shortparallel} + \cot\left(\frac{\gamma}{2}\right)\left(\bar{\xi}^{\shortparallel}\right)^2\right\}\mathbf{w}. \tag{A.32}$$

Equations (A.29) and (A.32) together define the relationship between the time derivative of the orientation vector and angular velocity,

$$\dot{\xi} = \left[\bar{\xi}\bar{\xi}^T - \frac{1}{2}\xi^{\shortparallel} - \frac{\|\xi\|}{2}\cot\left(\frac{\|\xi\|}{2}\right)\left(\bar{\xi}^{\shortparallel}\right)^2\right]\mathbf{w}. \tag{A.33}$$

The complexity of this expression clearly indicates the difficulty of implementing an extended Kalman filter with the angle-axis parameterization if temporal behavior is modelled directly in terms of angular velocity.

A.3 Quaternion Parameterization

In order to extend the operations of three-dimensional vector algebra to include multiplication and division, Hamilton (1843) introduced an algebra for four-dimensional numbers, or quaternions. The objective of this section is to provide an introduction to quaternion calculus with emphasis on their use in representing change-of-basis transformations. Results of the previous section on the Euler angle-axis parameterization will play a major role in the following derivations. Branets and Shmyglevski [83] present a very detailed treatment of the application of quaternions to rigid body rotation problems and provide an elegant interpretation of rotation in terms of spherical geometry. Hughes [84] treats several parameterizations, including quaternions, in applications to attitude dynamics.

A quaternion [29]-[37] [65]-[70], [83], [84], is a hypercomplex number which is expressed in terms of basis elements consisting of the real number $+1$, and three imaginary units i, j, k, which satisfy

$$\begin{aligned} i^2 = j^2 = k^2 = -1, \\ ij = -ji = k, \\ jk = -kj = i, \quad \text{and} \\ ki = -ik = j. \end{aligned} \tag{A.34}$$

A quaternion, $\mathbf{q}$ can be written[2] as

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = q_1 i + q_2 j + q_3 k + q_4. \tag{A.35}$$

Quaternions can be viewed as containing the real numbers $[0, 0, 0, a]^T$ with the real unit 1, the complex numbers $[b, 0, 0, a]^T$ with the two units 1 and i, and the vectors $[b, c, d, 0]^T$ in three-dimensional space. Some additional notation is required for developments of this section. Both quaternions and vectors are represented by bold face lower-case letters. As before, matrices are represented by bold upper-case letters (with the exception of matrix cross product operators). The vector part of a quaternion $\mathbf{q}$ is written as $\text{vect}(\mathbf{q}) = [q_1, q_2, q_3]^T$ while the real or scalar part of $\mathbf{q}$ is denoted

[2]Some authors place the scalar quaternion element, q_4, in the first position while others place the scalar element in the fourth position, as is the case here, when representing the quaternion in vector form.

by $\text{scal}(\mathbf{q}) = q_4$. With this notation,

$$\mathbf{q} = \text{vect}(\mathbf{q}) + \text{scal}(\mathbf{q}). \tag{A.36}$$

A quaternion formed from a vector $\mathbf{r} \in \mathbb{R}^3$ by adding a zero scalar part will be denoted by $\text{quat}(\mathbf{r})$.

Addition of quaternions and their multiplication by a scalar (which can also be viewed as a quaternion) are performed in the same way as in an ordinary vector space. Quaternion multiplication, which is denoted here with the symbol 'o', is defined in terms of the products of basis elements shown in (A.34). Quaternion multiplication is associative, and distributive with respect to addition, but it is noncommutative which is evident from the last three rows of (A.34), i.e. $ij \neq ji$ for example. For any quaternion $\mathbf{q}$, the conjugate, $\mathbf{q}^*$, of $\mathbf{q}$ is defined as

$$\mathbf{q}^* \triangleq -\text{vect}(\mathbf{q}) + \text{scal}(\mathbf{q}), \tag{A.37}$$

and the norm is defined as the Euclidean length of the 4-vector representing $\mathbf{q}$:

$$\|\mathbf{q}\| \triangleq \sqrt{q_1^2 + q_2^2 + q_3^2 + q_4^2}, \tag{A.38}$$

$$= \sqrt{\mathbf{q} \circ \mathbf{q}^*}, \tag{A.39}$$

where the second equation above follows from the rules of multiplication given in (A.34). Provided $\|\mathbf{q}\| \neq 0$, the inverse of $\mathbf{q}$ is

$$\mathbf{q}^{-1} \triangleq \frac{\mathbf{q}^*}{\|\mathbf{q}\|^2}, \tag{A.40}$$

and satisfies $\mathbf{q} \circ \mathbf{q}^{-1} = \mathbf{q}^{-1} \circ \mathbf{q} = 1$. For any quaternion $\mathbf{q}$, the corresponding unit quaternion $\bar{\mathbf{q}}$ is given by

$$\bar{\mathbf{q}} \triangleq \frac{\mathbf{q}}{\|\mathbf{q}\|}, \tag{A.41}$$

and note that $\bar{\mathbf{q}}^{-1} = \bar{\mathbf{q}}^*$. The set of unit quaternions can be placed in a two-to-one correspondence with elements of the three dimensional rotation matrices.

In Equation (A.13), the transformation from F_O to F_E was written as

$$\mathbf{I}_E^O = \mathbf{I}_3 + \sin(\gamma)\bar{\xi}^{\mathbf{I}} + [1 - \cos(\gamma)]\left(\bar{\xi}^{\mathbf{I}}\right)^2. \tag{A.42}$$

Using the familiar double-angle formulas

$$\sin(\gamma) \equiv 2\cos(\tfrac{\gamma}{2})\sin(\tfrac{\gamma}{2}), \text{ and} \tag{A.43}$$

$$\cos(\gamma) \equiv \cos^2(\tfrac{\gamma}{2}) - \sin^2(\tfrac{\gamma}{2}) \tag{A.44}$$

$$= 1 - 2\sin^2(\tfrac{\gamma}{2}) \tag{A.45}$$

in (A.42) gives

$$\mathbf{I}_E^O = \mathbf{I}_3 - 2\cos(\frac{\gamma}{2})\left[-\sin(\frac{\gamma}{2})\bar{\xi}\right]^{\natural} + 2\left(\left[[-\sin(\frac{\gamma}{2})\bar{\xi}\right]^{\natural}\right)^2. \tag{A.46}$$

The set of four parameters $[\eta^T, \varepsilon]^T$, where

$$\begin{aligned} \eta &\triangleq -\sin(\frac{\gamma}{2})\bar{\xi} \\ \varepsilon &\triangleq \cos(\frac{\gamma}{2}) \end{aligned} \tag{A.47}$$

are called *Euler Parameters* [84] and can be identified with the unit quaternion

$$\bar{\mathbf{q}} = [\eta^T, \varepsilon]^T. \tag{A.48}$$

With

$$\mathbf{q} = \|\mathbf{q}\|(\eta + \varepsilon), \tag{A.49}$$

substitution of (A.47) into (A.46) gives the following matrix representation for $\mathbf{I}_E^O$ in terms of the elements of a quaternion $\mathbf{q}$:

$$\mathbf{I}_E^O(\mathbf{q}) = \frac{1}{\|\mathbf{q}\|^2}\begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_1q_3 - q_2q_4) \\ 2(q_1q_2 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 + q_1q_4) \\ 2(q_1q_3 + q_2q_4) & 2(q_2q_3 - q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix}. \tag{A.50}$$

It can be verified by direct computation, using the multiplication rules in (A.34), that for any vector $\mathbf{r}$, the representation of $\mathbf{r}$ in F_E can be obtained from its representation in F_O according to

$$\begin{aligned} \mathbf{r}_E &= \mathbf{I}_E^O(\mathbf{q})\mathbf{r}_O \\ &= \text{vect}\left\{\mathbf{q}^{-1} \circ \text{quat}(\mathbf{r}_O) \circ \mathbf{q}\right\} \\ &= \text{vect}\left\{\bar{\mathbf{q}}^* \circ \text{quat}(\mathbf{r}_O) \circ \bar{\mathbf{q}}\right\} \end{aligned} \tag{A.51}$$

This property of quaternion calculus is extremely useful when multiple consecutive transformations (or rotations) must be performed as in the roll-pitch-yaw formulation which will be demonstrated in Section A.4.

Temporal behavior of the unit quaternion $\bar{\mathbf{q}}(t)$ can be obtained through differentiation of η and ε in (A.47) and the use of results in Equations (A.29) and (A.32) of Section 3.5.1. With some algebraic manipulation, one easily obtains

$$\dot{\eta} = -\frac{1}{2}\left(\eta^{\natural} + \varepsilon\mathbf{I}_3\right)\mathbf{w}, \text{ and} \tag{A.52}$$

$$\dot{\varepsilon} = -\frac{1}{2}\eta^T\mathbf{w}, \tag{A.53}$$

where $\mathbf{w} \equiv [\mathbf{w}_E^O]_E = [w_1, w_2, w_3]^T$ is, as before, the angular velocity vector of F_O with respect to F_E expressed in F_E. Two equivalent expressions for $\dot{\bar{\mathbf{q}}}(t)$ are then given by using (A.48) in (A.52) and (A.53):

$$\dot{\bar{\mathbf{q}}} = \frac{\partial \dot{\bar{\mathbf{q}}}}{\partial \mathbf{w}} \mathbf{w}, \text{ and} \tag{A.54}$$

$$\dot{\bar{\mathbf{q}}} = \mathbf{\Omega}[\mathbf{w}(t)]\bar{\mathbf{q}}(t), \tag{A.55}$$

where

$$\frac{\partial \dot{\bar{\mathbf{q}}}}{\partial \mathbf{w}} = -\frac{1}{2}\begin{bmatrix} \bar{q}_4 & -\bar{q}_3 & \bar{q}_2 \\ \bar{q}_3 & \bar{q}_4 & -\bar{q}_1 \\ -\bar{q}_2 & \bar{q}_1 & \bar{q}_4 \\ \bar{q}_1 & \bar{q}_2 & \bar{q}_3 \end{bmatrix}, \text{ and} \tag{A.56}$$

$$\mathbf{\Omega}[\mathbf{w}] = \frac{1}{2}\begin{bmatrix} 0 & -w_3 & w_2 & -w_1 \\ w_3 & 0 & -w_1 & -w_2 \\ -w_2 & w_1 & 0 & -w_3 \\ w_1 & w_2 & w_3 & 0 \end{bmatrix}. \tag{A.57}$$

Both (A.56) and (A.57) are required in the extended Kalman filtering equations.

With constant angular velocity, the solution of (A.55) with initial conditions given as $\bar{\mathbf{q}}(t_0)$ is simply

$$\bar{\mathbf{q}}(t) = \exp\left[\mathbf{\Omega} \cdot (t - t_0)\right] \bar{\mathbf{q}}(t_0). \tag{A.58}$$

Note that $\mathbf{\Omega}$ is skew symmetric and, since the matrix exponential of a skew symmetric is always orthogonal, time propagation of the quaternion in (A.58) with unit length is maintained. It can be verified by direct computation that

$$\mathbf{\Omega}^2 = -\frac{1}{4} \| \mathbf{w} \|^2 \mathbf{I}_4, \tag{A.59}$$

which implies that the highest power of $\mathbf{\Omega}$ in a power series is one. Expansion of (A.58) followed by repeated application of (A.59) gives

$$\bar{\mathbf{q}}(t) = \left[\cos\left(\|\mathbf{w}\|(t - t_0)/2\right) \mathbf{I}_4 + \frac{2}{\|\mathbf{w}\|} \sin\left(\|\mathbf{w}\|(t - t_0)/2\right) \mathbf{\Omega}\right] \bar{\mathbf{q}}(t_0). \tag{A.60}$$

Equation (A.60) is only valid when $\mathbf{w}$ is constant. Closed form solutions for $\bar{\mathbf{q}}(t)$ are also available [34] for rotation with constant precession.

A.4 Roll-Pitch-Yaw Parameterization

The roll-pitch-yaw representation [38, 63, 82] is a particular case of an Euler angle parameterization of the three-dimensional rotation group. The roll, pitch, and yaw

angles, denoted ϕ, θ, and ψ, respectively, define an ordered sequence of three plane rotations which can be used to express the basis transformation $\mathbf{I}_E^O$ as

$$\mathbf{I}_E^O = \exp(\psi \mathbf{e}_3^{\natural}) \exp(\theta \mathbf{e}_2^{\natural}) \exp(\phi \mathbf{e}_1^{\natural}), \tag{A.61}$$

where $\mathbf{e}_i$, $i = 1, 2, 3$, are the standard basis vectors of $\mathbb{R}^3$. Expanding the exponentials with (A.13) gives the three plane rotations,

$$\begin{aligned}
\exp(\phi \mathbf{e}_1^{\natural}) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi) & \cos(\phi) \end{bmatrix}, \\
\exp(\theta \mathbf{e}_2^{\natural}) &= \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix}, \text{ and} \\
\exp(\psi \mathbf{e}_3^{\natural}) &= \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\end{aligned} \tag{A.62}$$

The angles ϕ, θ, and ψ have positive sense defined by the right hand rule about their respective rotation axis. Specifically, the rotation from F_E to F_O is carried out by the following sequence:

1. rotate about $\mathbf{k}_E$ by the yaw angle ψ;
2. rotate about the axis $[-\sin(\psi), \cos(\psi), 0]_E^T$ by the pitch angle θ; and
3. rotate about the axis $[\cos(\psi)\cos(\theta), \sin(\psi)\cos(\theta), -\sin(\theta)]_E^T$ by the roll angle ϕ.

In order to avoid ambiguities in these angles their ranges are often limited to

$$\begin{aligned}
-\pi &\leq \psi < \pi, \\
-\tfrac{\pi}{2} &\leq \theta \leq \tfrac{\pi}{2}, \text{ and} \\
-\pi &\leq \phi < \pi.
\end{aligned} \tag{A.63}$$

The angular velocity of F_O with respect to F_E expressed in F_E, $\mathbf{w} \equiv [\mathbf{w}_E^O]_E$, is given in terms of the time derivatives of roll, pitch, and yaw angles as

$$[\mathbf{w}_E^O]_E = \mathbf{J}_{RPY} \begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix}, \tag{A.64}$$

where

$$\mathbf{J}_{RPY} = \begin{bmatrix} \cos(\psi)\cos(\theta) & -\sin(\psi) & 0 \\ \sin(\psi)\cos(\theta) & \cos(\psi) & 0 \\ -\sin(\theta) & 0 & 1 \end{bmatrix}. \tag{A.65}$$

contains the three rotational axis (from items 1,2, and 3 above) as its columns (which follows from the vector addition property of angular velocity). The mathematical singularities, $\theta = \pm\pi/2$, characterize the orientations where the determinant of the Jacobian matrix $\det[J_{RPY}] = -\cos(\theta)$ vanishes. At $\theta = \pi/2$, the roll and yaw angles are indistinguishable in the sense that the rotation matrix can be written as

$$\mathbf{I}_E^O(\theta = \frac{\pi}{2}) = \begin{bmatrix} 0 & \sin(\phi-\psi) & \cos(\phi-\psi) \\ 0 & \cos(\phi-\psi) & -\sin(\phi-\psi) \\ 1 & 0 & 0 \end{bmatrix}, \tag{A.66}$$

and hence depends only on $\theta = \pi/2$ and the difference $\phi - \psi$.

The sequence of roll, pitch and yaw transformations can be represented in terms of corresponding unit quaterions,

$$\bar{\mathbf{q}}_\phi \triangleq \begin{bmatrix} -\sin(\phi/2) \\ 0 \\ 0 \\ \cos(\phi/2) \end{bmatrix}, \quad \bar{\mathbf{q}}_\theta \triangleq \begin{bmatrix} 0 \\ -\sin(\theta/2) \\ 0 \\ \cos(\theta/2) \end{bmatrix}, \quad \bar{\mathbf{q}}_\psi \triangleq \begin{bmatrix} 0 \\ 0 \\ -\sin(\psi/2) \\ \cos(\theta/2) \end{bmatrix}, \tag{A.67}$$

according to

$$\begin{aligned} \mathbf{r}_E &= \mathbf{I}_E^O \mathbf{r}_O \\ &= \exp(\psi \mathbf{e}_3'') \left[\exp(\theta \mathbf{e}_2'') \left(\exp(\phi \mathbf{e}_1'') \mathbf{r}_O\right)\right] \\ &= \text{vect}\left\{\bar{\mathbf{q}}_\psi^* \circ \bar{\mathbf{q}}_\theta^* \circ \bar{\mathbf{q}}_\phi^* \circ \text{quat}(\mathbf{r}_O) \circ \bar{\mathbf{q}}_\phi \circ \bar{\mathbf{q}}_\theta \circ \bar{\mathbf{q}}_\psi\right\}. \end{aligned} \tag{A.68}$$

A single equivalent quaternion representing the basis transformation can be then defined as

$$\bar{\mathbf{q}}_{\text{equiv}} \triangleq \bar{\mathbf{q}}_\phi \circ \bar{\mathbf{q}}_\theta \circ \bar{\mathbf{q}}_\psi. \tag{A.69}$$

Noting that multiplication and conjugation rules for quaternions imply that

$$\bar{\mathbf{q}}_{\text{equiv}}^* = \bar{\mathbf{q}}_\psi^* \circ \bar{\mathbf{q}}_\theta^* \circ \bar{\mathbf{q}}_\phi^* \tag{A.70}$$

the transformation operation can be written as

$$\mathbf{I}_E^O \mathbf{r}_O = \text{vect}\left\{\bar{\mathbf{q}}_{\text{equiv}}^* \circ \text{quat}(\mathbf{r}_O) \circ \bar{\mathbf{q}}_{\text{equiv}}\right\}. \tag{A.71}$$

The Euler axis-angle representation corresponding to a particular set of roll, pitch and yaw angles can then easily be extracted from the equivalent quaternion.

By forming the product of the three plane rotations in (A.61) and defining the elements of $\mathbf{I}_E^O$ as ℓ_{ij}, the following algorithm can be used to extract a valid set of roll

(ϕ), pitch (θ), and yaw (ψ) angles:

$$
\begin{aligned}
&\text{Check for singularity with } \epsilon \ll 1 \\
&\text{if } ((|\ell_{11}| < \epsilon) \;\&\; (|\ell_{21}| < \epsilon)), \\
&\quad \phi = 0.0; \\
&\quad \psi = -\text{atan2}(\ell_{12}, \ell_{22}); \\
&\quad \text{if } (\ell_{31} > 0.0), \\
&\qquad \theta = \pi/2; \\
&\quad \text{else} \\
&\qquad \theta = -\pi/2; \\
&\text{else} \\
&\quad \phi = \text{atan2}(\ell_{32}, \ell_{33}); \\
&\quad \theta = \text{atan}\left(\frac{-\ell_{31}}{\sqrt{\ell_{11}^2 + \ell_{21}^2}}\right); \\
&\quad \psi = \text{atan2}(\ell_{21}, \ell_{11}); \\
&\text{end.}
\end{aligned}
\tag{A.72}
$$

Functions atan2(y, x) and atan(x) return, respectively, the four-quadrant arctangent of y/x and the two-quadrant arctangent of x, and are available in most programming languages. Note that in the singular orientations ($\theta = \pm\pi$), the roll angle is arbitrarily set to zero.

Appendix B
Kalman Filtering Review

This appendix provides a review and discussion of the linear discrete-time Kalman filter, the discrete-time extended Kalman filter, and the continuous-discrete Kalman filter. Local iterations which lead to the iterated extended Kalman filter are also introduced. Applications of Kalman filtering to trajectory estimation problems has received widespread attention since the introduction of Kalman's sequential state estimation technique [60]. Chang and Tabaczynski [62] give a thorough discussion of many important aspects of Kalman filtering approaches for target tracking problems. In general, the trajectory estimation problem is one of nonlinear estimation for which the optimal (conditional mean) nonlinear estimator cannot be realized with a finite-dimensional implementation. Furthermore, the mathematical model representing the target trajectory is seldom exact. Consequently, most practical implementations of nonlinear tracking filters are suboptimal.

The Extended Kalman Filter (EKF) is a suboptimal approach which applies linear Kalman filtering theory to a linear approximation of the nonlinear problem. A rigorous development of filtering equations requires the use of stochastic integral and differential equations and mathematical expectation conditioned on σ-algebras generated by the observation process, and formulates solutions through Ito calculus ([74, 81, for example]). In applications, however, the somewhat more intuitive approaches using "formal" manipulations of white noise processes and expectation conditioned on sets of observations are often preferred [62], [65], [23]–[44]. This approach will also be taken in the following review of the EKF formulation. The following subsections draw primarily from [74]-[78].

In many cases, the target tracking problem can be formulated in terms of discrete-time system (plant) and measurement models. In the case of image-based tracking, measurements are taken from image frames which arrive at discrete instants in time. The measurement model for this class of problems, therefore, is always discrete. In addition, the perspective projection image formation model results in a nonlinear measurement model. If the system model is continuous in time, it can often be discretized, or the "continuous-discrete" [74] form of the filter can be employed.

Notation and filtering equations for the discrete-time linear state estimation prob-

lem are reviewed in Section B.1, with extensions to approximate filters for the nonlinear discrete-time problem in Section B.2 and the continuous-discrete form of extended Kalman filtering in Section B.3. Finally, Section B.4 describes local iterations which enhance filter performance in the presence of nonlinear measurement models.

B.1 Discrete-Time Kalman Filter

Consider a linear discrete-time dynamical system model given by

$$\mathbf{x}(k+1) = \mathbf{\Phi}(k)\mathbf{x}(k) + \mathbf{B}(k)\mathbf{u}(k) + \mathbf{G}(k)\mathbf{w}(k), \tag{B.1}$$

where $k \sim$ an integer valued sample index, time $t = kT$ where T is the sample period, and

$$\begin{aligned}
\mathbf{x} \in \mathbb{R}^n &\sim \text{vector of state variables,} \\
\mathbf{u} \in \mathbb{R}^q &\sim \text{deterministic input vector,} \\
\mathbf{w} \in \mathbb{R}^p &\sim \text{zero mean, white, Gaussian, system noise} \\
&\quad \text{process with covariance matrix } \mathbf{Q(k)}, \\
\mathbf{\Phi} \in \mathbb{R}^{n\times n} &\sim \text{state transition matrix,} \\
\mathbf{B} \in \mathbb{R}^{n\times q} &\sim \text{input matrix mapping the input vector} \\
&\quad \mathbf{u} \text{ into the state space, and} \\
\mathbf{G} \in \mathbb{R}^{n\times p} &\sim \text{matrix mapping the noise process} \\
&\quad \mathbf{w} \text{ into the state space.}
\end{aligned} \tag{B.2}$$

The linear, discrete-time measurement model is given by

$$\mathbf{z}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{v}(k), \tag{B.3}$$

where

$$\begin{aligned}
\mathbf{z} \in \mathbb{R}^m &\sim \text{vector of observations or measurements,} \\
\mathbf{H} \in \mathbb{R}^{n\times m} &\sim \text{measurement matrix relating state variables} \\
&\quad \text{to measurement variables, and} \\
\mathbf{v} \in \mathbb{R}^m &\sim \text{zero mean, white, Gaussian, measurement noise} \\
&\quad \text{process with covariance matrix } \mathbf{R}(k).
\end{aligned} \tag{B.4}$$

With definitions (B.2) and (B.4), the system and measurement noise processes satisfy

$$\begin{aligned}
\mathcal{E}\{\mathbf{w}(k)\} = \mathbf{0} \qquad & \mathcal{E}\left\{\mathbf{w}(k)\mathbf{w}^T(j)\right\} = \mathbf{Q}(k)\delta_{kj} \\
\mathcal{E}\{\mathbf{v}(k)\} = \mathbf{0} \qquad & \mathcal{E}\left\{\mathbf{v}(k)\mathbf{v}^T(j)\right\} = \mathbf{R}(k)\delta_{kj},
\end{aligned} \tag{B.5}$$

where $\mathcal{E}\{\cdot\}$ denotes expectation and δ_{jk} is the Kronecker delta product. In the present analysis, $\mathbf{w}$ and $\mathbf{v}$ are also assumed to be uncorrelated, i.e.,

$$\mathcal{E}\left\{\mathbf{w}(k)\mathbf{v}^T(j)\right\} = \mathbf{0}, \tag{B.6}$$

for all integers j and k.

The expectation of the state $\mathbf{x}$ at time step k conditioned on all measurements up to time step j is denoted by

$$\hat{\mathbf{x}}(k|j) \triangleq \mathcal{E}\{\mathbf{x}(k)|Z_j\} \tag{B.7}$$

where $Z_j = \{\mathbf{z}(i),\ i = 1, 2, \cdots, j\}$ represents the set of all measurements up to and including the jth time step. It is precisely the conditional mean in (B.7), usually with $j = k$, which the optimal filter is required to estimate. In the case of linear system and measurement models, assuming exact mathematical models and characterization of the noise processes, the Kalman filter provides a recursive "formula" for this optimal estimate. It is well known that in this case the Kalman filter provides an unbiased, minimum variance, consistent estimate $\hat{\mathbf{x}}(k|k)$, which is also the maximum likelihood estimate of the true state $\mathbf{x}(k)$.

The error between the filter state and the true state is given by

$$\tilde{\mathbf{x}}(k|j) \triangleq \mathbf{x}(k) - \hat{\mathbf{x}}(k|j), \tag{B.8}$$

with covariance matrix

$$\mathbf{P}(k|j) \triangleq \mathcal{E}\left\{\tilde{\mathbf{x}}(k|j)\tilde{\mathbf{x}}^T(k|j)\right\}. \tag{B.9}$$

The development of the filtering equations assumes that the true initial state of the system $\mathbf{x}(0) \sim \mathrm{N}(\bar{\mathbf{x}}_0, \mathbf{P}_0)$[1]. Throughout this work, it is assumed that an initial estimate, $\hat{\mathbf{x}}(0|-1)$, of the mean of the true initial state, $\mathcal{E}\{\mathbf{x}(0)\}$, is available just prior to taking initial measurements, $\mathbf{z}(0)$. In theory, then, the recursive filter is initialized with

$$\begin{aligned} \hat{\mathbf{x}}(0|-1) &= \mathcal{E}\{\mathbf{x}(0)\} \triangleq \bar{\mathbf{x}}_0, \text{ and} \\ \mathbf{P}(0|-1) &= \mathcal{E}\left\{\tilde{\mathbf{x}}(0|-1)\tilde{\mathbf{x}}^T(0|-1)\right\} \triangleq \mathbf{P}_0. \end{aligned} \tag{B.10}$$

In practice, however, the filter is often initialized with an unbiased estimator of $\bar{\mathbf{x}}_0$ for which some *a priori* statistical information is known of the form

$$\hat{\mathbf{x}}(0|-1) \sim \mathrm{N}(\bar{\mathbf{x}}_0, \mathbf{P}_0). \tag{B.11}$$

After $k-1$ measurements have been taken, the recursive equations of the linear Kalman filter are summarized as follows:

1. State and covariance prediction (or time update) -

$$\begin{aligned} \hat{\mathbf{x}}(k|k-1) &= \mathbf{\Phi}(k-1)\hat{\mathbf{x}}(k-1|k-1) + \mathbf{B}(k-1)\mathbf{u}(k-1), \text{ and} \\ \mathbf{P}(k|k-1) &= \mathbf{\Phi}(k-1)\mathbf{P}(k-1|k-1)\mathbf{\Phi}^T(k-1) \\ &\quad + \mathbf{G}(k-1)\mathbf{Q}(k-1)\mathbf{G}^T(k-1). \end{aligned} \tag{B.12}$$

[1]The notation $\mathrm{N}(\mathbf{m}, \mathbf{C})$ denotes a normally distributed random variable with mean $\mathbf{m}$ and covariance $\mathbf{C}$.

2. Measurement update computed following the kth measurement -

$$\begin{aligned} \mathbf{K}(k) &= \mathbf{P}(k|k-1)\mathbf{H}^T(k) \\ &\quad \cdot \left[\mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}^T(k) + \mathbf{R}(k)\right]^{-1}, \qquad \text{(B.13)} \\ \hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\left[\mathbf{z}(k) - \mathbf{H}(k)\hat{\mathbf{x}}(k|k-1)\right], \text{ and} \\ \mathbf{P}(k|k) &= \left[\mathbf{I} - \mathbf{K}(k)\mathbf{H}(k)\right]\mathbf{P}(k|k-1). \end{aligned}$$

An equivalent form, known as the "information form", of the measurement update stage of the filter, which is often used in place of (B.13), is given by

$$\begin{aligned} \mathbf{P}^{-1}(k|k) &= \mathbf{P}^{-1}(k|k-1) + \mathbf{H}^T(k)\mathbf{R}^{-1}(k)\mathbf{H}(k), \\ \mathbf{K}(k) &= \mathbf{P}(k|k)\mathbf{H}^T(k)\mathbf{R}^{-1}(k), \text{ and} \qquad \text{(B.14)} \\ \hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\left[\mathbf{z}(k) - \mathbf{H}(k)\hat{\mathbf{x}}(k|k-1)\right]. \end{aligned}$$

In any implementation of the above filtering equations, the following parameters must be specified in addition to the measurement sequence: 1) an initial state estimate $\hat{\mathbf{x}}(0|-1)$; 2) an estimate of the initial error covariance $\mathbf{P}_0$; 3) the system or plant noise covariances $\mathbf{Q}(k)$; and 4) the measurement or observation noise covariances $\mathbf{R}(k)$. The initial state estimate and its error covariance matrix often result from a batch estimation technique applied to the first few image frames [30, 32, 38]. In simulations, the measurement noise covariance is known, however, in applications to real imagery, determination of the actual measurement noise is difficult since this involves modeling of various sources of error in the imaging system and feature detection algorithms. Specification of the system noise covariance $\mathbf{Q}(k)$ is seldom exact in object tracking problems since unmodeled dynamics resulting from manoeuvring object trajectories are difficult to quantify. As a result, $\mathbf{Q}(k)$ together with the matrix $\mathbf{G}(k)$ are often treated as "tuning" parameters which are selected to give satisfactory convergent behavior of the filter over a wide range of applications.

In Equation (B.12) the previous state estimate $\hat{\mathbf{x}}(k-1|k-1)$ is propagated in time between measurements according to the noise-free dynamic system model to obtain a state estimate $\hat{\mathbf{x}}(k|k-1)$ just prior to a measurement event. Once measurements are available, the state estimate is corrected in (B.13) by an amount proportional to the current error $[\mathbf{z}(k) - \mathbf{H}(k)\hat{\mathbf{x}}(k|k-1)]$ called the innovation. The proportionality factor $\mathbf{K}(k)$ is called the Kalman gain which, from (B.14), is proportional to the uncertainty in the state estimate $\mathbf{P}(k|k)$ and inversely proportional to the measurement uncertainty $\mathbf{R}(k)$. The system noise covariance $\mathbf{Q}(k)$ which is mapped to the state space through $\mathbf{G}(k)$ adds a positive definite matrix to the filter error covariance matrix during each time update (B.12) which reflects a "diffusion" of confidence in the state estimate over the time interval between measurements. In the absence of system noise ($\mathbf{Q} \equiv \mathbf{0}$) the error covariance and gain matrices may decreases monotonically in the presence of measurement data, depending on the state transition matrix through the similarity transformation in (B.12), i.e., consider (B.14) for the scalar case with

$\mathbf{H} \equiv 1$. In this case, the filter eventually neglects further measurements and operates solely on the system model which, with $\mathbf{Q} \equiv \mathbf{0}$, is assumed to be exact.

B.2 Discrete-Time Extended Kalman Filter

The problem of estimating the state of a nonlinear system can be formulated as an extension of mean square estimation to dynamic systems. Formally, the method for deriving the fundamental equations of the optimal filter is quite similar to that used to design the optimal linear filter (Kalman filter) for linear systems with Gaussian noise processes. However, in the nonlinear case, the explicit analytical solution cannot, in general, be found and hence approximate numerical solutions are often considered.

In the case of linear systems, all of the variables and processes involved (states, noises, measurements, estimates, and errors) are usually assumed to be normally distributed. Therefore, mean values and covariance matrices are "sufficient statistics." In the nonlinear case, on the other hand, the state, observation, and estimation processes are not, in general, normally distributed even if the input noise sources are. As a result, the entire probability distribution may be required for optimal filter design. In addition, the filter performance and parameters are often dependent on the filter estimate so that it is even more difficult to obtain bounds for filter performance.

Consider the nonlinear system model given by

$$\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k)] + \mathbf{B}(k)\mathbf{u}(k) + \mathbf{g}[\mathbf{x}(k)]\mathbf{w}(k), \tag{B.15}$$

and the nonlinear measurement model

$$\mathbf{z}(k) = \mathbf{h}[\mathbf{x}(k)] + \mathbf{v}(k), \tag{B.16}$$

where $\mathbf{f} : R^n \mapsto R^n$, $\mathbf{g} : R^n \mapsto (R^{n \times p})$, and $\mathbf{h} : R^n \mapsto R^m$ are, in general, nonlinear functions of the state vector $\mathbf{x}(k)$. The mapping of the input vector $\mathbf{u}$ to the state space may also be nonlinear, but this case will not be considered here and, in any event, does not alter the filter structure significantly.

The Extended Kalman Filter (EKF) is based on the concept that if the nonlinearities in $\mathbf{f}$ and $\mathbf{h}$ are sufficiently smooth, these functions can be expanded in a Taylor series and approximated by linear terms. Furthermore, if the filter works properly, it will provide a reasonable estimate of the conditional mean $\mathcal{E}\{\mathbf{x}(k)|Z_k\}$ which is what the optimal filter is required to estimate. For this reason, the Taylor series expansion of nonlinear terms is formulated about the most recent state estimate. The primary consequence of this "feedback" of the estimate into the model is that the EKF is a nonlinear estimation technique.

With the definitions

$$\mathbf{\Phi}(k) \triangleq \left.\frac{\partial \mathbf{f}[\mathbf{x}]}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}(k|k)}, \tag{B.17}$$

$$\mathbf{G}(k) \triangleq \mathbf{g}[\hat{\mathbf{x}}(k|k)], \quad \text{and} \tag{B.18}$$

$$\mathbf{H}(k) \triangleq \left.\frac{\partial \mathbf{h}[\mathbf{x}]}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}(k|k-1)}, \tag{B.19}$$

and neglecting all terms beyond first order in the Taylor series expansions, approximate linear system and measurement models can be written, respectively, as

$$\mathbf{x}(k+1) = \mathbf{\Phi}(k)\mathbf{x}(k) + \tilde{\mathbf{u}}(k) + \mathbf{G}(k)\mathbf{w}(k), \text{ and} \tag{B.20}$$

$$\mathbf{z}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{y}(k) + \mathbf{v}(k), \tag{B.21}$$

where

$$\tilde{\mathbf{u}}(k) \triangleq \mathbf{f}[\hat{\mathbf{x}}(k|k)] - \mathbf{\Phi}(k)\hat{\mathbf{x}}(k|k) + \mathbf{B}(k)\mathbf{u}(k), \text{ and} \tag{B.22}$$

$$\mathbf{y}(k) \triangleq \mathbf{h}[\hat{\mathbf{x}}(k|k-1)] - \mathbf{H}(k)\hat{\mathbf{x}}(k|k-1). \tag{B.23}$$

The EKF equations are formulated, essentially, as the standard Kalman filtering equations for the linear system given by (B.17)-(B.23).

After $k-1$ measurements have been taken, the recursive equations of the extended Kalman filter are summarized as follows:

1. State and covariance prediction (or time update) -

$$\begin{aligned} \hat{\mathbf{x}}(k|k-1) &= \mathbf{f}[\hat{\mathbf{x}}(k-1|k-1)] + \mathbf{B}(k)\mathbf{u}(k), \text{ and} \\ \hat{\mathbf{P}}(k|k-1) &= \mathbf{\Phi}(k-1)\hat{\mathbf{P}}(k-1|k-1)\mathbf{\Phi}^T(k-1) \\ &\quad + \mathbf{G}(k-1)\mathbf{Q}(k-1)\mathbf{G}^T(k-1). \end{aligned} \tag{B.24}$$

2. Measurement update computed following the kth measurement -

$$\begin{aligned} \mathbf{K}(k) &= \hat{\mathbf{P}}(k|k-1)\mathbf{H}^T(k) \\ &\quad \cdot \left[\mathbf{H}(k)\hat{\mathbf{P}}(k|k-1)\mathbf{H}^T(k) + \mathbf{R}(k)\right]^{-1}, \\ \hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\left\{\mathbf{z}(k) - \mathbf{h}[\hat{\mathbf{x}}(k|k-1)]\right\}, \text{ and} \\ \hat{\mathbf{P}}(k|k) &= [\mathbf{I} - \mathbf{K}(k)\mathbf{H}(k)]\,\hat{\mathbf{P}}(k|k-1). \end{aligned} \tag{B.25}$$

The accent on $\hat{\mathbf{P}}$ has been used to emphasize the fact that the covariance estimate is dependent on the filter estimate.

B.3 Continuous-Discrete Extended Kalman Filter

In Section B.2, both the system and measurement models were assumed to be available in discrete-time form. In general, however, the system model is often developed in continuous-time form, while measurements are taken at discrete instants in time.

Consider a continuous-time system which is described by the stochastic differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}, t) + \mathbf{G}(t)\mathbf{w}(t), \quad t \geq 0, \quad \mathbf{x}(0) \sim N(\bar{\mathbf{x}}_0, \mathbf{P}_0), \tag{B.26}$$

where $\mathbf{w}(t)$ is a zero-mean white Gaussian noise process with

$$\mathcal{E}\left\{\mathbf{w}(t)\mathbf{w}^T(\tau)\right\} = \mathbf{Q}(t)\delta(t-\tau), \tag{B.27}$$

and $\delta(t-\tau)$ is the Dirac-delta function. The system model may also depend on a deterministic input vector $\mathbf{u}(t)$. To simplify the notation, this dependence is not shown explicitly here and, in any event, does not alter the filter structure significantly. As in the discrete-time case of Section B.2, the nonlinear system model of (B.26) is linearized approximately through a first-order Taylor series expansion about a nominal trajectory $\hat{\mathbf{x}}(t)$ which satisfies

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\hat{\mathbf{x}}, t), \quad t \geq 0. \tag{B.28}$$

The state perturbation is defined as

$$\tilde{\mathbf{x}}(t) \triangleq \mathbf{x}(t) - \hat{\mathbf{x}}(t), \tag{B.29}$$

which satisfies the stochastic differential equation

$$\begin{aligned} \dot{\tilde{\mathbf{x}}}(t) &= \mathbf{f}(\mathbf{x}, t) - \mathbf{f}(\hat{\mathbf{x}}, t) + \mathbf{G}(t)\mathbf{w}(t) \\ &\approx \mathbf{F}[t; \hat{\mathbf{x}}]\tilde{\mathbf{x}}(t) + \mathbf{G}(t)\mathbf{w}(t), \end{aligned} \tag{B.30}$$

with the initial condition $\tilde{\mathbf{x}}(0) \sim N(\bar{\mathbf{x}}_0 - \hat{\mathbf{x}}(0), \mathbf{P}_0)$ and where

$$\mathbf{F}[t; \hat{\mathbf{x}}] \triangleq \left.\frac{\partial \mathbf{f}(\mathbf{x}, t)}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}}. \tag{B.31}$$

It should be noted that $\mathbf{F}$ is evaluated at particular values of $\hat{\mathbf{x}}(t)$ and hence is a function of t only. Specifically, as in the discrete-time case of Section B.2, the nominal trajectory is taken as the most recent state estimate which, for the linearized system model, is given by

$$\hat{\mathbf{x}}(t_k) \equiv \hat{\mathbf{x}}(k|k). \tag{B.32}$$

In principle, equation (B.30) can be written in discrete-time form as

$$\tilde{\mathbf{x}}(k+1) = \hat{\mathbf{\Phi}}(t_{k+1}, t_k)\tilde{\mathbf{x}}(k) + \mathbf{w}_d(k+1), \tag{B.33}$$

where $\hat{\mathbf{\Phi}}(t_{k+1}, t_k)$ is the state transition matrix which satisfies [74]

$$\begin{aligned} \frac{\mathrm{d}\hat{\mathbf{\Phi}}(t,\tau)}{\mathrm{d}t} &= \mathbf{F}[t; \mathbf{x}(t)]\hat{\mathbf{\Phi}}(t,\tau), \\ \hat{\mathbf{\Phi}}(\tau,\tau) &= \mathbf{I}_n, \quad \forall \tau, \\ \hat{\mathbf{\Phi}}(t,\tau)\hat{\mathbf{\Phi}}(\tau,\zeta) &= \hat{\mathbf{\Phi}}(t,\zeta), \quad \forall t, \tau, \zeta, \end{aligned} \tag{B.34}$$

and $\mathbf{w}_d(k+1)$ is a zero-mean, white Gaussian sequence, $\mathbf{w}_d(k+1) \sim N(0, \mathbf{Q}_d(k+1))$, where

$$\mathbf{Q}_d(k+1) = \int_{t_k}^{t_{k+1}} \hat{\mathbf{\Phi}}(t_{k+1}, \tau)\mathbf{G}(\tau)\mathbf{Q}(\tau)\mathbf{G}^T(\tau)\hat{\mathbf{\Phi}}^T(t_{k+1}, \tau)\, d\tau. \tag{B.35}$$

If $\mathbf{F}$ is approximately constant over the interval $(t_k, t_{k+1}]$, where $t_{k+1} - t_k = T_k$, the state transition matrix for each sample interval can be estimated as

$$\hat{\mathbf{\Phi}}(t_{k+1}, t_k) = \exp(\mathbf{F}T_k). \tag{B.36}$$

In general, however, the computation of the state transition matrix is often difficult and, therefore, replaced by numerical integration of the Riccati covariance equation [74]

$$\begin{aligned} \dot{\hat{\mathbf{P}}}(t|t_k) &= \mathbf{F}[t; \hat{\mathbf{x}}(t|t_k)]\hat{\mathbf{P}}(t|t_k) \\ &\quad + \hat{\mathbf{P}}(t|t_k)\mathbf{F}^T[t; \hat{\mathbf{x}}(t|t_k)] + \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^T(t). \end{aligned} \tag{B.37}$$

The nonlinear, discrete-time measurement model is again given by (B.16) and is treated in a manner identical to that used in Section B.2. The extended Kalman filter is then formulated based on the approximate linear perturbation equation (B.30), or its discrete form (B.33), and the linear approximation of the measurement model (B.21). However, in this case, the state prediction is computed by integration of the nonlinear noise-free system differential equation. After $k-1$ measurements have been taken, the recursive equations of the continuous-discrete extended Kalman filter are summarized as follows:

1. State and covariance prediction (or time update) - Either

$$\begin{aligned} \hat{\mathbf{x}}(k|k-1) &= \hat{\mathbf{x}}(k-1|k-1) + \int_{t_{k-1}}^{t_k} \mathbf{f}[\hat{\mathbf{x}}(t|t_{k-1}), t]\, dt \text{ and} \\ \hat{\mathbf{P}}(k|k-1) &= \hat{\mathbf{P}}(k-1|k-1) + \int_{t_{k-1}}^{t_k} \dot{\hat{\mathbf{P}}}[t; \hat{\mathbf{x}}(t|t_{k-1})]\, dt, \end{aligned} \tag{B.38}$$

where $\dot{\hat{\mathbf{P}}}[t; \mathbf{x}(t|t_{k-1})]$ is defined in (B.37), or equivalently,

$$\begin{aligned} \hat{\mathbf{x}}(k|k-1) &= \hat{\mathbf{\Phi}}(k-1)\hat{\mathbf{x}}(k-1|k-1) \\ \hat{\mathbf{P}}(k|k-1) &= \hat{\mathbf{\Phi}}(k-1)\hat{\mathbf{P}}(k-1|k-1)\hat{\mathbf{\Phi}}^T(k-1) + \mathbf{Q}_d(k). \end{aligned} \tag{B.39}$$

2. Measurement update computed following the kth measurement -

$$\begin{aligned} \mathbf{K}(k) &= \hat{\mathbf{P}}(k|k-1)\mathbf{H}^T(k) \\ &\quad \cdot \left[\mathbf{H}(k)\hat{\mathbf{P}}(k|k-1)\mathbf{H}^T(k) + \mathbf{R}(k)\right]^{-1}, \\ \hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\left\{\mathbf{z}(k) - \mathbf{h}[\hat{\mathbf{x}}(k|k-1)]\right\}, \text{ and} \\ \hat{\mathbf{P}}(k|k) &= [\mathbf{I} - \mathbf{K}(k)\mathbf{H}(k)]\,\hat{\mathbf{P}}(k|k-1). \end{aligned} \tag{B.40}$$

Again, the accent on $\hat{\mathbf{P}}$ has been used to emphasize the fact that the covariance estimate is dependent on the filter estimate.

B.4 Local Iterations

The Iterated Extended Kalman Filter (IEKF) is designed to reduce the effect of measurement nonlinearities on the performance of the linear filter applied to a linearized nonlinear system [74, p. 279]. Between observations the conditional mean and covariance matrix propagate according to the equations of (B.24). Following the time update of the state and covariance from time step $k-1$ to time step k, and with the kth measurement available, the iteration is initialized with

$$\hat{\mathbf{x}}(k|k)_0 = \hat{\mathbf{x}}(k|k-1). \tag{B.41}$$

Iterations are carried out on the measurement update equations in (B.25) according to the following algorithm:

1. Linearize the measurement equations about the current iterate -

$$\mathbf{H}(k)_n = \left.\frac{\partial \mathbf{h}[\mathbf{x}]}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}(k|k)_n}, \tag{B.42}$$

2. Compute the Kalman gain -

$$\begin{aligned}\mathbf{K}(k)_n &= \hat{\mathbf{P}}(k|k-1)\mathbf{H}^T(k)_n \\ &\quad \cdot \left[\mathbf{H}(k)_n \hat{\mathbf{P}}(k|k-1)\mathbf{H}^T(k)_n + \mathbf{R}(k)\right]^{-1},\end{aligned} \tag{B.43}$$

3. Compute the measurement update -

$$\begin{aligned}\hat{\mathbf{x}}(k|k)_{n+1} &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)_n \left\{\mathbf{z}(k) - \mathbf{h}[\hat{\mathbf{x}}(k|k)_n]\right. \\ &\quad \left. - \mathbf{H}(k)_n \left[\hat{\mathbf{x}}(k|k-1) - \hat{\mathbf{x}}(k|k)_n\right]\right\}.\end{aligned} \tag{B.44}$$

Iteration is terminated when there is no significant change in consecutive iterates. Then the last iterate, $\hat{\mathbf{x}}(k|k)_N$ say, is taken for the estimate $\hat{\mathbf{x}}(k|k)$ and the covariance matrix is computed based on the linearization about $\hat{\mathbf{x}}(k|k)_N$,

$$\hat{\mathbf{P}}(k|k) = \left[\mathbf{I} - \mathbf{K}(k)_N \mathbf{H}(k)_N\right] \hat{\mathbf{P}}(k|k-1). \tag{B.45}$$

The iterative procedure described by (B.42)-(B.44) can be interpreted as a modified Newton-Raphson search for the conditional mode of the posterior density of the state estimate given the measurement at time step k, assuming the prior density is Gaussian [74, pp. 349-351]. The mode of the posterior density is then used for the conditional mean. In nonlinear problems, these local iterations can be expected to produce a biased estimate since the mode is not, in general, equivalent to the mean. However, as the error variance becomes small, so does the bias in this estimate.

Appendix C
Cramer-Rao Covariance Bounds

The extended Kalman filter is generally a suboptimal state estimation technique. A natural consideration in any estimation problem is how well the proposed technique performs when compared to the optimal performance possible. A powerful result which provides a performance assessment of parameter estimation techniques is the Cramer-Rao inequality [79]. The following derivation draws from methods outlined in [61, 62, 74, 79, 80].

Consider a nonlinear discrete-time estimation problem with noise-free system model

$$\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k)] \tag{C.1}$$

and measurement model

$$\mathbf{z}(k) = \mathbf{h}[\mathbf{x}(k)] + \mathbf{v}(k), \tag{C.2}$$

where $\mathbf{v}(k) \sim \mathrm{N}(\mathbf{0}, \mathbf{R})$ and is temporally white. Assume further that some unbiased *a priori* statistical information about $\mathbf{x}(0)$ is available of the form

$$\mathbf{x}_0 \sim \mathrm{N}(\mathbf{x}(0), \mathbf{P}_0), \tag{C.3}$$

where $\mathbf{x}_0$ and $\mathbf{v}(k)$ are independent. The objective is to determine how well $\mathbf{x}(N)$, $N \geq 0$, can be estimated given the sequence of observations Z_N^*, where

$$Z_N^* \triangleq \{\mathbf{x}_0, \mathbf{z}(0), \mathbf{z}(1), \cdots, \mathbf{z}(N)\}. \tag{C.4}$$

Under the stated assumptions of independence and normal distributions, the joint conditional probability density function of Z_N^* given $\mathbf{x}(N)$ can be written, from (C.2) and (C.3), as

$$p_N \triangleq p_{Z_N^*|\mathbf{x}(N)} = \frac{1}{A} e^{-\frac{1}{2}\Psi[Z_N^*; \mathbf{x}(N)]}, \tag{C.5}$$

where[1]

$$A \triangleq (2\pi)^{n/2}|\mathbf{P}_0|^{1/2} \prod_{k=0}^{N} [(2\pi)^{m/2}|\mathbf{R}(k)|^{1/2}], \tag{C.6}$$

[1]The notation $|\mathbf{P}|$ denotes the determinant of $\mathbf{P}$.

$$\Psi[Z_N^*; \mathbf{x}(N)] \triangleq (\mathbf{x}_0 - \bar{\mathbf{x}}_0)^T \mathbf{P}_0^{-1} (\mathbf{x}_0 - \bar{\mathbf{x}}_0) + \sum_{k=0}^{N} [\mathbf{z}(k) - \bar{\mathbf{z}}(k)]^T \mathbf{R}^{-1}(k) [\mathbf{z}(k) - \bar{\mathbf{z}}(k)], \quad \text{(C.7)}$$

and

$$\bar{\mathbf{x}}_0 \triangleq \mathcal{E}\{\mathbf{x}_0\} = \mathbf{x}(0), \text{ and} \quad \text{(C.8)}$$

$$\bar{\mathbf{z}}(k) \triangleq \mathcal{E}\{\mathbf{z}(k)\} = \mathbf{h}[\mathbf{x}(k)]. \quad \text{(C.9)}$$

The Cramer-Rao lower bound for the error covariance matrix $\mathbf{S}(N)$ of an unbiased estimator $\hat{\mathbf{x}}(N)$ of the state $\mathbf{x}(N)$, where the estimate is based on the observation sequence Z_N^*, is given by

$$\mathbf{S}(N) \triangleq \mathcal{E}\{[\hat{\mathbf{x}}(N) - \mathbf{x}(N)][\hat{\mathbf{x}}(N) - \mathbf{x}(N)]^T\} \geq \mathbf{J}^{-1}(N), \quad \text{(C.10)}$$

where $\mathbf{J}(N)$ is Fisher's information matrix. The matrix inequality $\mathbf{S} \geq \mathbf{J}^{-1}$ is equivalent to stating that $(\mathbf{S} - \mathbf{J}^{-1})$ is positive semi-definite. In particular, since the diagonal elements of a positive semi-definite matrix are non-negative, the diagonal elements of $\mathbf{J}^{-1}(N)$ provide the estimation error variance lower bounds for the corresponding elements of $\hat{\mathbf{x}}(N)$. In the linear system and measurement case, (B.1) and (B.3), the discrete Kalman filter given by (B.12) and (B.13) provides the minimum variance unbiased estimate. If the system noise $\mathbf{w}(k)$ in (B.1) is zero, then for the linear case

$$\mathbf{S}(N) = \mathbf{P}(N|N) = \mathbf{J}^{-1}(N). \quad \text{(C.11)}$$

In the nonlinear EKF formulation, $\mathbf{S}(N)$ and $\hat{\mathbf{P}}(N|N)$ are not, in general, equivalent.

Fisher's information measure on $\mathbf{x}(N)$ contained in Z_N^* is given by two equivalent representations [80, pp. 91-93]:

$$\mathbf{J}(N) = \mathcal{E}\left\{ \left(\frac{\partial \ln p_N}{\partial \mathbf{x}(N)} \right)^T \left(\frac{\partial \ln p_N}{\partial \mathbf{x}(N)} \right) \middle| \mathbf{x}(N) \right\}, \text{ or} \quad \text{(C.12)}$$

$$\mathbf{J}(N) = -\mathcal{E}\left\{ \left(\frac{\partial^2 \ln p_N}{\partial \mathbf{x}(N)^2} \right) \middle| \mathbf{x}(N) \right\}, \quad \text{(C.13)}$$

where the gradient $\partial \ln p_N / \partial \mathbf{x}(N)$ is a row vector. The inequality (C.10) with $\mathbf{J}(N)$ expressed as in (C.12) is termed the "Cramer-Rao inequality."

Taking the logarithm of p_N in (C.5) yields

$$\ln p_N = -\ln A - \frac{1}{2} \Psi[Z_N^*; \mathbf{x}(N)]. \quad \text{(C.14)}$$

The gradient of (C.14) is given by the row vector

$$\frac{\partial \ln p_N}{\partial \mathbf{x}(N)} = (\mathbf{x}_0 - \bar{\mathbf{x}}_0)^T \mathbf{P}_0^{-1} \frac{\partial \bar{\mathbf{x}}_0}{\partial \mathbf{x}(N)} + \sum_{k=0}^{N} [\mathbf{z}(k) - \bar{\mathbf{z}}(k)]^T \mathbf{R}^{-1}(k) \frac{\partial \bar{\mathbf{z}}(k)}{\partial \mathbf{x}(N)}. \quad \text{(C.15)}$$

Appendix C
Cramer-Rao Covariance Bounds

The extended Kalman filter is generally a suboptimal state estimation technique. A natural consideration in any estimation problem is how well the proposed technique performs when compared to the optimal performance possible. A powerful result which provides a performance assessment of parameter estimation techniques is the Cramer-Rao inequality [79]. The following derivation draws from methods outlined in [61, 62, 74, 79, 80].

Consider a nonlinear discrete-time estimation problem with noise-free system model

$$\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k)] \tag{C.1}$$

and measurement model

$$\mathbf{z}(k) = \mathbf{h}[\mathbf{x}(k)] + \mathbf{v}(k), \tag{C.2}$$

where $\mathbf{v}(k) \sim \mathrm{N}(\mathbf{0}, \mathbf{R})$ and is temporally white. Assume further that some unbiased *a priori* statistical information about $\mathbf{x}(0)$ is available of the form

$$\mathbf{x}_0 \sim \mathrm{N}(\mathbf{x}(0), \mathbf{P}_0), \tag{C.3}$$

where $\mathbf{x}_0$ and $\mathbf{v}(k)$ are independent. The objective is to determine how well $\mathbf{x}(N)$, $N \geq 0$, can be estimated given the sequence of observations Z_N^*, where

$$Z_N^* \triangleq \{\mathbf{x}_0, \mathbf{z}(0), \mathbf{z}(1), \cdots, \mathbf{z}(N)\}. \tag{C.4}$$

Under the stated assumptions of independence and normal distributions, the joint conditional probability density function of Z_N^* given $\mathbf{x}(N)$ can be written, from (C.2) and (C.3), as

$$p_N \triangleq p_{Z_N^*|\mathbf{x}(N)} = \frac{1}{A} e^{-\frac{1}{2}\Psi[Z_N^*; \mathbf{x}(N)]}, \tag{C.5}$$

where[1]

$$A \triangleq (2\pi)^{n/2} |\mathbf{P}_0|^{1/2} \prod_{k=0}^{N} [(2\pi)^{m/2} |\mathbf{R}(k)|^{1/2}], \tag{C.6}$$

[1]The notation $|\mathbf{P}|$ denotes the determinant of $\mathbf{P}$.

$$\Psi[Z_N^*; \mathbf{x}(N)] \triangleq (\mathbf{x}_0 - \bar{\mathbf{x}}_0)^T \mathbf{P}_0^{-1} (\mathbf{x}_0 - \bar{\mathbf{x}}_0) + \sum_{k=0}^{N} [\mathbf{z}(k) - \bar{\mathbf{z}}(k)]^T \mathbf{R}^{-1}(k) [\mathbf{z}(k) - \bar{\mathbf{z}}(k)], \quad \text{(C.7)}$$

and

$$\bar{\mathbf{x}}_0 \triangleq \mathcal{E}\{\mathbf{x}_0\} = \mathbf{x}(0), \text{ and} \quad \text{(C.8)}$$

$$\bar{\mathbf{z}}(k) \triangleq \mathcal{E}\{\mathbf{z}(k)\} = \mathbf{h}[\mathbf{x}(k)]. \quad \text{(C.9)}$$

The Cramer-Rao lower bound for the error covariance matrix $\mathbf{S}(N)$ of an unbiased estimator $\hat{\mathbf{x}}(N)$ of the state $\mathbf{x}(N)$, where the estimate is based on the observation sequence Z_N^*, is given by

$$\mathbf{S}(N) \triangleq \mathcal{E}\{[\hat{\mathbf{x}}(N) - \mathbf{x}(N)][\hat{\mathbf{x}}(N) - \mathbf{x}(N)]^T\} \geq \mathbf{J}^{-1}(N), \quad \text{(C.10)}$$

where $\mathbf{J}(N)$ is Fisher's information matrix. The matrix inequality $\mathbf{S} \geq \mathbf{J}^{-1}$ is equivalent to stating that $(\mathbf{S} - \mathbf{J}^{-1})$ is positive semi-definite. In particular, since the diagonal elements of a positive semi-definite matrix are non-negative, the diagonal elements of $\mathbf{J}^{-1}(N)$ provide the estimation error variance lower bounds for the corresponding elements of $\hat{\mathbf{x}}(N)$. In the linear system and measurement case, (B.1) and (B.3), the discrete Kalman filter given by (B.12) and (B.13) provides the minimum variance unbiased estimate. If the system noise $\mathbf{w}(k)$ in (B.1) is zero, then for the linear case

$$\mathbf{S}(N) = \mathbf{P}(N|N) = \mathbf{J}^{-1}(N). \quad \text{(C.11)}$$

In the nonlinear EKF formulation, $\mathbf{S}(N)$ and $\hat{\mathbf{P}}(N|N)$ are not, in general, equivalent.

Fisher's information measure on $\mathbf{x}(N)$ contained in Z_N^* is given by two equivalent representations [80, pp. 91-93]:

$$\mathbf{J}(N) = \mathcal{E}\left\{ \left(\frac{\partial \ln p_N}{\partial \mathbf{x}(N)} \right)^T \left(\frac{\partial \ln p_N}{\partial \mathbf{x}(N)} \right) \middle| \mathbf{x}(N) \right\}, \text{ or} \quad \text{(C.12)}$$

$$\mathbf{J}(N) = -\mathcal{E}\left\{ \left(\frac{\partial^2 \ln p_N}{\partial \mathbf{x}(N)^2} \right) \middle| \mathbf{x}(N) \right\}, \quad \text{(C.13)}$$

where the gradient $\partial \ln p_N / \partial \mathbf{x}(N)$ is a row vector. The inequality (C.10) with $\mathbf{J}(N)$ expressed as in (C.12) is termed the "Cramer-Rao inequality."

Taking the logarithm of p_N in (C.5) yields

$$\ln p_N = -\ln A - \frac{1}{2} \Psi[Z_N^*; \mathbf{x}(N)]. \quad \text{(C.14)}$$

The gradient of (C.14) is given by the row vector

$$\frac{\partial \ln p_N}{\partial \mathbf{x}(N)} = (\mathbf{x}_0 - \bar{\mathbf{x}}_0)^T \mathbf{P}_0^{-1} \frac{\partial \bar{\mathbf{x}}_0}{\partial \mathbf{x}(N)} + \sum_{k=0}^{N} [\mathbf{z}(k) - \bar{\mathbf{z}}(k)]^T \mathbf{R}^{-1}(k) \frac{\partial \bar{\mathbf{z}}(k)}{\partial \mathbf{x}(N)}. \quad \text{(C.15)}$$

To evaluate (C.12), consider the four terms of the outer product of (C.15):

$$\begin{aligned}
&\left(\frac{\partial p_N}{\partial \mathbf{x}(N)}\right)^T \left(\frac{\partial p_N}{\partial \mathbf{x}(N)}\right) = \\
&\quad \left(\frac{\partial \bar{\mathbf{x}}_0}{\partial \mathbf{x}(N)}\right)^T \mathbf{P}_0^{-1}(\mathbf{x}_0 - \bar{\mathbf{x}}_0)(\mathbf{x}_0 - \bar{\mathbf{x}}_0)^T \mathbf{P}_0^{-1} \left(\frac{\partial \bar{\mathbf{x}}_0}{\partial \mathbf{x}(N)}\right) \\
&\quad + \sum_{k=0}^{N} \left(\frac{\partial \bar{\mathbf{x}}_0}{\partial \mathbf{x}(N)}\right)^T \mathbf{P}_0^{-1}(\mathbf{x}_0 - \bar{\mathbf{x}}_0)\,[\mathbf{z}(k) - \bar{\mathbf{z}}(k)]^T\, \mathbf{R}^{-1}(k) \left(\frac{\partial \bar{\mathbf{z}}(k)}{\partial \mathbf{x}(N)}\right) \\
&\quad + \sum_{k=0}^{N} \left(\frac{\partial \bar{\mathbf{z}}(k)}{\partial \mathbf{x}(N)}\right)^T \mathbf{R}^{-1}(k)\,[\mathbf{z}(k) - \bar{\mathbf{z}}(k)]\,(\mathbf{x}_0 - \bar{\mathbf{x}}_0)^T \mathbf{P}_0^{-1} \left(\frac{\partial \bar{\mathbf{x}}_0}{\partial \mathbf{x}(N)}\right) \\
&\quad + \left\{\sum_{k=0}^{N} \left(\frac{\partial \bar{\mathbf{z}}(k)}{\partial \mathbf{x}(N)}\right)^T \mathbf{R}^{-1}(k)\,[\mathbf{z}(k) - \bar{\mathbf{z}}(k)]\right\} \\
&\qquad \cdot \left\{\sum_{j=0}^{N} [\mathbf{z}(j) - \hat{\mathbf{z}}(j)]^T\, \mathbf{R}^{-1}(j) \left(\frac{\partial \bar{\mathbf{z}}(j)}{\partial \mathbf{x}(N)}\right)\right\}.
\end{aligned} \tag{C.16}$$

The conditional expectation in (C.12) can be computed from (C.16) with the following observations:

1. The Jacobian matrices $\partial \bar{\mathbf{x}}_0 / \partial \mathbf{x}(N)$ and $\partial \bar{\mathbf{z}}(k) / \partial \mathbf{x}(N)$ are deterministic;

2. From (C.2), (C.8), and (C.9), the random variables $(\mathbf{x}_0 - \bar{\mathbf{x}}_0)$ and $[\mathbf{z}(k) - \bar{\mathbf{z}}(k)]$, $k = 0, 1, \cdots, N$, are independent with zero mean and hence the second and third terms of (C.16) do not contribute to the expectation and the product of summations in the last term of (C.16) reduces to a single summation of products with $k = j$;

3. By definitions (C.2) and (C.3),

$$\begin{aligned}
\mathcal{E}\left\{(\mathbf{x}_0 - \bar{\mathbf{x}}_0)(\mathbf{x}_0 - \bar{\mathbf{x}}_0)^T\right\} &= \mathbf{P}_0 \text{ and} \\
\mathcal{E}\left\{[\mathbf{z}(k) - \bar{\mathbf{z}}(k)][\mathbf{z}(k) - \bar{\mathbf{z}}(k)]^T\right\} &= \mathbf{R}(k).
\end{aligned}$$

Furthermore, from the definitions given in (C.8) and (C.9),

$$\frac{\partial \bar{\mathbf{x}}_0}{\partial \mathbf{x}(N)} = \frac{\partial \mathbf{x}(0)}{\partial \mathbf{x}(N)} \tag{C.17}$$

and

$$\begin{aligned}
\frac{\partial \bar{\mathbf{z}}(k)}{\partial \mathbf{x}(N)} &= \frac{\partial \mathbf{h}[\mathbf{x}(k)]}{\partial \mathbf{x}(N)} \\
&= \mathbf{H}_t(k) \frac{\partial \mathbf{x}(k)}{\partial \mathbf{x}(N)},
\end{aligned} \tag{C.18}$$

where

$$\mathbf{H}_t(k) \triangleq \left.\frac{\partial \mathbf{h}[\mathbf{x}]}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}(k)}. \tag{C.19}$$

The subscript "t" on $\mathbf{H}_t$ is used to emphasize that (C.19) is evaluated along the true trajectory $\mathbf{x}(k)$ rather than along the estimated trajectory $\hat{\mathbf{x}}(k)$ as in the usual EKF realization.

Performing the expectation of (C.16) in accordance with (C.12) and substituting (C.17), (C.18), and (C.19) yields Fisher's information matrix:

$$\begin{aligned}\mathbf{J}(N) &= \left(\frac{\partial \mathbf{x}(0)}{\partial \mathbf{x}(N)}\right)^T \mathbf{P}_0^{-1} \left(\frac{\partial \mathbf{x}(0)}{\partial \mathbf{x}(N)}\right) \\ &\quad + \sum_{k=0}^{N} \left(\frac{\partial \mathbf{x}(k)}{\partial \mathbf{x}(N)}\right)^T \mathbf{H}_t^T(k)\mathbf{R}^{-1}(k)\mathbf{H}_t(k) \left(\frac{\partial \mathbf{x}(k)}{\partial \mathbf{x}(N)}\right).\end{aligned} \tag{C.20}$$

Evaluation of the alternate representation for Fisher's information matrix in (C.13) is most easily accomplished through component-wise differentiation of (C.14). Consider a vector $\xi \in R^n$ and a scalar-valued function $\mathcal{F}(\xi)$ given by

$$\mathcal{F}(\xi) = \frac{1}{2}\xi^T \Gamma \xi, \tag{C.21}$$

where $\Gamma \in R^{n\times n}$ is positive definite symmetric. Equation (C.21) represents the general form of $\ln p_N$ in (C.14) which must be differentiated. Provided the components of ξ admit second derivatives with respect to components of a vector $\vartheta \in R^n$,

$$\begin{aligned}\left[\frac{\partial^2 \mathcal{F}(\xi)}{\partial \vartheta^2}\right]_{i,j} &= \frac{\partial^2 \mathcal{F}(\xi)}{\partial \vartheta_i \partial \vartheta_j} \\ &= \sum_{\ell,k} \left[\frac{\partial \xi_\ell}{\partial \vartheta_i}\Gamma_{\ell,k}\frac{\partial \xi_k}{\partial \vartheta_j} + \xi_\ell \Gamma_{\ell,k}\frac{\partial^2 \xi_k}{\partial \vartheta_i \partial \vartheta_j}\right] \\ &= \left[\left(\frac{\partial \xi}{\partial \vartheta}\right)^T \Gamma \left(\frac{\partial \xi}{\partial \vartheta}\right)\right]_{i,j} + \sum_{\ell,k}\left[\xi_\ell \Gamma_{\ell,k}\frac{\partial^2 \xi_k}{\partial \vartheta_i \partial \vartheta_j}\right].\end{aligned} \tag{C.22}$$

Taking the expectation of (C.22) in accordance with (C.13) in a sequence of steps with $\vartheta = \mathbf{x}(N)$ and first with $\xi = (\mathbf{x}_0 - \bar{\mathbf{x}}_0)$ and then with $\xi = [\mathbf{z}(k) - \bar{\mathbf{z}}(k)]$, $k = 0, 1, 2, \ldots, N$, which implies that in all steps ξ has zero mean and deterministic first and second partial derivatives, eliminates the second term on the right of (C.22) giving, with (C.17), (C.18), and (C.19), the previous result (C.20) as expected.

Finally, from (C.1)

$$\frac{\partial \mathbf{x}(k+1)}{\partial \mathbf{x}(N)} = \Phi_t(k)\frac{\partial \mathbf{x}(k)}{\partial \mathbf{x}(N)}, \quad \text{for } 0 \le k \le N-1, \tag{C.23}$$

where

$$\mathbf{\Phi}_t(k) \triangleq \left.\frac{\partial \mathbf{f}[\mathbf{x}]}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}(k)}, \quad \text{for } 0 \leq k \leq N-1 \tag{C.24}$$

is evaluated along the true trajectory as shown. In particular, with $\mathbf{I}_n$ representing the $n \times n$ identity matrix,

$$\frac{\partial \mathbf{x}(N)}{\partial \mathbf{x}(N)} = \mathbf{I}_n = \left[\downarrow \prod_{i=N-1}^{k} \mathbf{\Phi}_t(i)\right] \frac{\partial \mathbf{x}(k)}{\partial \mathbf{x}(N)}, \tag{C.25}$$

where the notation "$\downarrow\prod$" indicates that the index i decreases from left to right in the matrix product. The Jacobian matrices in (C.20) are then computed as

$$\frac{\partial \mathbf{x}(k)}{\partial \mathbf{x}(N)} = \prod_{i=k}^{N-1} \mathbf{\Phi}_t^{-1}(i). \tag{C.26}$$

Substitution of (C.26) into (C.20) gives a closed-form expression for Fisher's information matrix for the system and measurement models defined in (C.1) and (C.2):

$$\begin{aligned} \mathbf{J}(N) \;=\; & \left(\prod_{i=0}^{N-1} \mathbf{\Phi}_t^{-1}(i)\right)^T \mathbf{P}_0^{-1} \left(\prod_{i=0}^{N-1} \mathbf{\Phi}_t^{-1}(i)\right) \\ & + \sum_{k=0}^{N} \left(\prod_{i=k}^{N-1} \mathbf{\Phi}_t^{-1}(i)\right)^T \mathbf{H}_t^T(k)\mathbf{R}^{-1}(k)\mathbf{H}_t(k) \left(\prod_{i=k}^{N-1} \mathbf{\Phi}_t^{-1}(i)\right). \end{aligned} \tag{C.27}$$

A more useful form for practical computation is a recursive formulation which follows from (C.27) by inspection:

1. Initialization –

$$\mathbf{J}(0) = \mathbf{P}_0^{-1} + \mathbf{H}_t^T(0)\mathbf{R}^{-1}(0)\mathbf{H}_t(0); \tag{C.28}$$

2. Recursive evaluation for $k \geq 0$ –

$$\mathbf{J}(k+1) = \left[\mathbf{\Phi}_t^{-1}(k)\right]^T \mathbf{J}(k) \left[\mathbf{\Phi}_t^{-1}(k)\right] + \mathbf{H}_t^T(k)\mathbf{R}^{-1}(k)\mathbf{H}_t(k). \tag{C.29}$$

If the information matrix $\mathbf{J}(k)$ remains singular in the absence of prior information, i.e. $\mathbf{P}_0^{-1} = 0$, then the system has an unobservable subspace. The converse is not generally true, that is the bounds may decrease even though the system is unobservable. Observability of nonlinear systems, although often difficult to establish, can be investigated locally through methods from differential geometry [72].

The above derivation has been given for discrete-time, nonlinear system and measurement models. Furthermore, this derivation assumes an unbiased estimation procedure. A more general result for an estimator with state dependent bias $\mathbf{b}(\mathbf{x})$ is given by [78]

$$\mathbf{S}(N) \geq \left(\mathbf{I}_n + \frac{\partial \mathbf{b}(\mathbf{x})}{\partial \mathbf{x}}\right) \mathbf{J}^{-1} \left(\mathbf{I}_n + \frac{\partial \mathbf{b}(\mathbf{x})}{\partial \mathbf{x}}\right)^T, \tag{C.30}$$

where $\mathbf{S}(N)$ is defined in (C.10) and the Jacobian of $\mathbf{b}$ is evaluated along the true trajectory. However, the bias $\mathbf{b}(\mathbf{x})$ is generally not available in analytical form. Unfortunately, a state dependent bias can result in higher or lower bounds in comparison to $\mathbf{J}^{-1}$: for example, consider $\mathbf{b}(\mathbf{x}) = -\mathbf{x}/2$ which gives $\mathbf{S} \geq \mathbf{J}^{-1}/4$, while $\mathbf{b}(\mathbf{x}) = \mathbf{x}$ gives $\mathbf{S} \geq 4\mathbf{J}^{-1}$. As a result, most analysis proceed under the assumption of unbiased estimation resulting in approximate covariance bounds.

Broida and Chellappa [28, 30, 31, 37], Young and Chellappa [34], Taylor [61], and Chang [62] all employ Cramer-Rao analysis for filter performance evaluation. Broida and Chellappa, and Young and Chellappa do not consider *a priori* information. Taylor does consider information provided by the initial state estimate and treats the continuous-time system case. In the continuous-discrete filtering techniques of Section B.3, results in [61] indicate that approximate covariance bounds can be computed as in (C.28) and (C.29), but with $\mathbf{\Phi}_t(k)$ replaced with $\hat{\mathbf{\Phi}}_t(k, k-1)$ defined in (2.11).

The Cramer Rao bounds shown in Section 4.3 do not decrease monotonically. This seems to contradict the notion that as more measurements become available, more and more information should be accumulated about the observed process and, as a result, the CRLB's would be expected to decrease. Fisher's information matrix defined in (C.12) provides a measure of information contained in Z_N^* concerning the state $\mathbf{x}(N)$. As time, N, increases, the measurement set Z_N^* accumulates more measurements but, in addition, the state $\mathbf{x}(N)$ may vary with time. As discussed by Young and Chellappa [34], the noise-free dynamical model implies a one-to-one deterministic relationship between $\mathbf{x}(0)$ and $\mathbf{x}(N)$ so that

$$p_{Z_N^*|\mathbf{x}(N)} = p_{Z_N^*|\mathbf{x}(0)}, \tag{C.31}$$

and

$$\frac{\partial p_{Z_N^*|\mathbf{x}(N)}}{\partial \mathbf{x}(N)} = \left[\frac{\partial p_{Z_N^*|\mathbf{x}(0)}}{\partial \mathbf{x}(0)}\right] \frac{\partial \mathbf{x}(0)}{\partial \mathbf{x}(N)}. \tag{C.32}$$

Inversion of Fisher's information matrix then gives

$$\mathbf{J}^{-1}(N) = \left(\frac{\partial \mathbf{x}(0)}{\partial \mathbf{x}(N)}\right)^{-1} \mathbf{K}^{-1}(N) \left(\frac{\partial \mathbf{x}(0)}{\partial \mathbf{x}(N)}\right)^{-T}, \tag{C.33}$$

where $(\cdot)^{-T}$ denotes the inverse transpose and

$$\begin{aligned} \mathbf{K}^{-1}(N) &\triangleq \left[\mathcal{E}\left\{\left(\frac{\partial \ln p_{Z_N^*|\mathbf{x}(0)}}{\partial \mathbf{x}(N)}\right)^T \left(\frac{\partial \ln p_{Z_N^*|\mathbf{x}(0)}}{\partial \mathbf{x}(N)}\right) \middle| \mathbf{x}(N)\right\},\right]^{-1} \\ &= \left[\mathbf{P}_0^{-1} + \sum_{n=0}^{N} \left(\frac{\partial \mathbf{x}(0)}{\partial \mathbf{x}(N)}\right)^T \mathbf{H}_t^T(n)\mathbf{R}^{-1}(n)\mathbf{H}_t(n) \left(\frac{\partial \mathbf{x}(0)}{\partial \mathbf{x}(N)}\right)\right]^{-1}. \end{aligned} \tag{C.34}$$

Cramer Rao bounds for estimation of the initial state $\mathbf{x}(0)$ from the measurement set Z_N^* are the diagonal elements of $\mathbf{K}^{-1}(N)$, and these decrease monotonically. The bounds on $\mathbf{x}(N)$, however, depend on the time-varying congruent transformation of $\mathbf{K}^{-1}(N)$ in (C.34) and need not decrease monatonically.

Appendix D
Measurement Model Jacobian

Implementation of the extended Kalman filter and estimation of Cramer Rao bounds requires the computation of $\mathbf{r}^i_{C_j}$, $\mathbf{I}^O_E(\zeta)$ and $\mathbf{h}[\mathbf{x}]$, which are all components of the measurement model, as well as the Jacobian of $\mathbf{h}[\mathbf{x}]$,

$$\mathbf{H}[\mathbf{x}] \triangleq \frac{\partial \mathbf{h}[\mathbf{x}]}{\partial \mathbf{x}}, \tag{D.1}$$

from the most recent state estimate. Computation of $\mathbf{H}[\mathbf{x}]$ for the three parameterizations of rotational motion is a tedious exercise in differentiation, but, with algebraic manipulation and simplification, can be written in simple form which admits computationally efficient implementation.

The Jacobian matrix, $\mathbf{H} \in R^{2N_f N_C \times N}$, can be written as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_{N_C} \end{bmatrix} \tag{D.2}$$

where $\mathbf{H}_j$ is the approximate linear measurement matrix corresponding to the jth camera. Each $\mathbf{H}_j$ can be further partitioned according to the feature point of interest as

$$\mathbf{H}_j = \begin{bmatrix} \mathbf{H}^1_j \\ \mathbf{H}^2_j \\ \vdots \\ \mathbf{H}^{N_f}_j \end{bmatrix}, \tag{D.3}$$

$$= \begin{bmatrix} \mathbf{P}^1_j & \mathbf{0} & \mathbf{W}^1_j & \mathbf{0} & \mathbf{S}^1_j & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{P}^2_j & \mathbf{0} & \mathbf{W}^2_j & \mathbf{0} & \mathbf{0} & \mathbf{S}^2_j & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}^{N_f}_j & \mathbf{0} & \mathbf{W}^{N_f}_j & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{S}^{N_f}_j \end{bmatrix}, \tag{D.4}$$

where each $\mathbf{H}^i_j \in \mathbb{R}^{2\times N}$, $i = 1, 2, \ldots, N_f$, $j = 1, 2, \ldots, N_C$ is defined as

$$\mathbf{H}^i_j \triangleq \frac{\partial \mathbf{h}^i_j[\mathbf{x}]}{\partial \mathbf{x}}. \tag{D.5}$$

The submatrices $\mathbf{P}_j^i \in \mathbb{R}^{2\times 3}$, $\mathbf{W}_j^i \in \mathbb{R}^{2\times(3 \text{ or } 4)}$, depending on the parameterization of rotational motion, and $\mathbf{S}_j^i \in \mathbb{R}^{2\times 3}$ can be written, after some algebraic manipulation, as follows:

$$\begin{aligned}
\mathbf{P}_j^i &\triangleq \frac{\partial \mathbf{h}_j^i[\mathbf{x}]}{\partial \mathbf{p}} \\
&= \frac{L_j}{\left(r_{C_j,1}^i\right)^2}\begin{bmatrix} -(\mathbf{r}_{C_j}^i)^T\mathbf{k}_E^{\mathbb{I}} \\ (\mathbf{r}_{C_j}^i)^T\mathbf{j}_E^{\mathbb{I}} \end{bmatrix}\mathbf{I}_{C_j}^E, \qquad \text{(D.6)} \\
\mathbf{W}_j^i &\triangleq \frac{\partial \mathbf{h}_j^i[\mathbf{x}]}{\partial \zeta} \\
&= \mathbf{P}_j^i\frac{\partial \mathbf{I}_E^O(\zeta)}{\partial \zeta}\mathbf{r}_O^i, \qquad \text{(D.7)} \\
\mathbf{S}_j^i &\triangleq \frac{\partial \mathbf{h}_j^i[\mathbf{x}]}{\partial \mathbf{r}_O^i} \\
&= \mathbf{P}_j^i\mathbf{I}_E^O. \qquad \text{(D.8)}
\end{aligned}$$

All entries of $\mathbf{H}$ other than those specified above are zero. In the above equations, $\mathbf{j}_E = [0,1,0]^T$, $\mathbf{k}_E = [0,0,1]^T$, and $\mathbf{j}_E^{\mathbb{I}}$ and $\mathbf{k}_E^{\mathbb{I}}$ are the matrix cross product operators associated with $\mathbf{j}_E$ and $\mathbf{k}_E$, respectively. In Equation (D.7), the notation $\partial \mathbf{I}_E^O/\partial\zeta$ denotes a three dimensional entity which premultiplies a vector $\vartheta \in \mathbb{R}^3$ to yield a matrix according to the rule

$$\left[\frac{\partial \mathbf{I}_E^O(\zeta)}{\partial \zeta}\vartheta\right]_{ij} = \sum_{k=1}^{3}\frac{\partial [\mathbf{I}_E^O]_{ik}}{\partial \zeta_j}\vartheta_k. \qquad \text{(D.9)}$$

The general form of the measurement model is common to all parameterizations of rotational motion, except that different parameterizations will yield different expressions for $\partial \mathbf{I}_E^O(\zeta)/\partial\zeta$. These expressions are now stated for each parameterization.

Angle-Axis Parameterization: The measurement model Jacobian can be computed with

$$\frac{\partial \mathbf{I}_E^O(\xi)}{\partial \xi_j} = \frac{\{1-\cos(\gamma)\}}{\gamma}\beta_j + \cos(\gamma)\bar{\xi}_j\bar{\xi}^{\mathbb{I}} + \frac{\sin(\gamma)}{\gamma}\Gamma_j + \sin(\gamma)\bar{\xi}_j\left(\bar{\xi}^{\mathbb{I}}\right)^2, \qquad \text{(D.10)}$$

where

$$\begin{aligned}
\gamma &\triangleq \|\xi\|, \qquad \text{(D.11)} \\
\bar{\xi} &\triangleq \frac{\xi}{\|\xi\|}, \qquad \text{(D.12)} \\
\Gamma_j &\triangleq \mathbf{e}_j^{\mathbb{I}} - \bar{\xi}_j\bar{\xi}^{\mathbb{I}}, \qquad \text{(D.13)} \\
\beta_j &\triangleq \Gamma_j\bar{\xi}^{\mathbb{I}} + \bar{\xi}^{\mathbb{I}}\Gamma_j, \qquad \text{(D.14)}
\end{aligned}$$

and $\bar{\xi}_j$ denotes the jth component of the unit vector $\bar{\xi}$, and $\mathbf{e}_j$ denotes the jth standard basis vector of $\mathbb{R}^3$. For very small γ, it is apparent from (D.10)-(D.14) that

$$\mathbf{I}_E^O(\xi) \sim \mathbf{I}_3 + \xi^\natural, \text{ and} \tag{D.15}$$

$$\frac{\partial \mathbf{I}_E^O(\xi)}{\partial \xi_j} \sim \frac{\gamma}{2}\beta + e_j^\natural. \tag{D.16}$$

This information can be used to eliminate numerical difficulties encountered in applications when (D.10) must be computed with $\gamma = \|\xi\| \sim 0$.

Quaternion Parameterization: The Jacobian of the measurement model is computed with

$$\frac{\partial \mathbf{I}_E^O(\mathbf{q})}{\partial q_1} = 2\begin{bmatrix} q_1 & q_2 & q_3 \\ q_2 & -q_1 & q_4 \\ q_3 & -q_4 & -q_1 \end{bmatrix}, \tag{D.17}$$

$$\frac{\partial \mathbf{I}_E^O(\mathbf{q})}{\partial q_2} = 2\begin{bmatrix} -q_2 & q_1 & -q_4 \\ q_1 & q_2 & q_3 \\ q_4 & q_3 & -q_2 \end{bmatrix}, \tag{D.18}$$

$$\frac{\partial \mathbf{I}_E^O(\mathbf{q})}{\partial q_3} = 2\begin{bmatrix} -q_3 & q_4 & q_1 \\ -q_4 & -q_3 & q_2 \\ q_1 & q_2 & q_3 \end{bmatrix}, \text{ and} \tag{D.19}$$

$$\frac{\partial \mathbf{I}_E^O(\mathbf{q})}{\partial q_4} = 2\begin{bmatrix} q_4 & q_3 & -q_2 \\ -q_3 & q_4 & q_1 \\ q_2 & -q_1 & q_4 \end{bmatrix}. \tag{D.20}$$

Roll-Pitch-Yaw Parameterization: The measurement model Jacobian is computed with

$$\frac{\partial \mathbf{I}_E^O(\zeta)}{\partial \phi} = \mathbf{I}_E^O \mathbf{e}_1^\natural, \tag{D.21}$$

$$\frac{\partial \mathbf{I}_E^O(\zeta)}{\partial \theta} = \exp(\psi \mathbf{e}_3^\natural)\mathbf{e}_2^\natural \exp(\theta \mathbf{e}_2^\natural)\exp(\phi \mathbf{e}_1^\natural), \text{ and} \tag{D.22}$$

$$\frac{\partial \mathbf{I}_E^O(\zeta)}{\partial \psi} = \mathbf{e}_3^\natural \mathbf{I}_E^O, \tag{D.23}$$

where, again, the $\mathbf{e}_j$, $j = 1, 2, 3$, are the standard basis vectors of $\mathbb{R}^3$.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

Defence Research Establishment Suffield
Box 4000, Medicine Hat, AB T1A 8K6

2. SECURITY CLASSIFICATION (overall security classification of the document including special warning terms if applicable)

Unclassified

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title.)

Motion and Structure Estimation of Manoeuvring Objects in Multiple-Camera Image Sequences

4. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.)

Aitken, Victor C.

5. DATE OF PUBLICATION (month and year of publication of document)

November 1992

6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)

144

6b. NO. OF REFS (total cited in document)

84

6. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Research Report

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

Defence Research Establishment Suffield
Box 4000, Medicine Hat, AB T1A 8K6

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

SR 577

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(x) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availabilty (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both offical languages unless the text is bilingual).

Estimation of structure and six-degree-of-freedom motion of manoeuvring objects through measurements of feature positions in long, multiple-camera image sequences is widely recognized to have broad industrial, military and space applications, particularly in the control of autonomous systems. This report focuses on the case of manoeuvring objectives thereby removing restrictive assumptions concerning the mode of translational and rotational motion which are commonly employed in many existing methods. Object manoeuvres, being "smooth" and time correlated, are modelled as first-order Gauss-Markov processes for both translational and rotational motion. In the literature, rotational motion is often parameterized with unit quaternions even though constrains on the quaternion norm are not easily enforced, roll-pitch-yaw parameterizations have been reported to be poorly behaved and have led to computationally demanding implementations, and results using the Euler angle-axis parameterization in recursive motion and structure estimation are not available. In this report, six-degree-of-freedom, nonlinear, approximate state estimation filters for quaternion, roll-pitch-yaw and angle-axis parameterizations are compared in terms of estimation performance for manoeuvring object trajectories. Special consideration is given to the problem of imposing unit norm on the estimated quaternion since previously proposed methods led to filter instability, particularly in angular velocity estimation. Methods proposed herein not only demonstrate how previous quaternion-based algorithms might be extended to track manoeuvring objects observed in multiple-camera image sequences, but also provide two much simpler alternatives using angle-axis and roll-pitch-yaw parameterizations. The angle-axis filter was found to give the best overall performance with a significant reduction in computational requirements in comparison to the quaternion filter.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Motion
Motion Analysis
Structure from Motion
Manoeuvring Objects
Target Tracking
Kalman Filtering
Extended Kalman Filtering
Quaternions
Angle Axis
Roll Pitch Yaw
Rotational Motion
Stereo
Image Sequences